

2 32  
GTT TTA CAC CAT CAC CAA AAC CCA GAA TTT TAT GAT GAG ATT AAA ATA GAG TTG CCC ACT  
val leu his his his gln asn pro glu phe tyr asp glu ile lys ile glu leu pro thr

62 92  
CAG CTG CAT GAA AAG CAC CAC CTG TTG CTC ACA TTC TTC CAT GTC AGC TGT GAC AAC TCA  
gln leu his glu lys his his leu leu leu thr phe phe his val ser cys asp asn ser

122 152  
AGT AAA GGA AGC ACG AAG AAG AGG GAT GTC GTT GAA ACC CAA GTT GGC TAC TCC TGG CTT  
ser lys gly ser thr lys lys arg asp val val glu thr gln val gly tyr ser trp leu

182 212  
CCC CTC CTG AAA GAC GGA AGG GTG GTG ACA AGC GAG CAG CAC ATC CCG GTC TCG GCG AAC  
pro leu leu lys asp gly arg val val thr ser glu gln his ile pro val ser ala asn

242 272  
CTT CCT TCG GGC TAT CTT GGC TAC CAA GAG CTT GGG ATG GGC AGG CAT TAT GGT CCG GAA  
leu pro ser gly tyr leu gly tyr gln glu leu gly met gly arg his tyr gly pro glu

302 332  
ATT AAA TGG GTA GAT GGA GGC AAG CCA CTG CTG AAA ATT TCC ACT CAT CTG GTT TCT ACA  
ile lys trp val asp gly gly lys pro leu leu lys ile ser thr his leu val ser thr

362 392  
GTG TAT ACT CAG GAT CAG CAT TTA CAT AAT TTT TTC CAG TAC TGT CAG AAA ACC GAA TCT  
val tyr thr gln asp gln his leu his asn phe phe gln tyr cys gln lys thr glu ser

422 452  
GGA GCC CAA GCC TTA GGA AAC GAA CTT GTA AAG TAC CTT AAG AGT CTG CAT GCG ATG GAA  
gly ala gln ala leu gly asn glu leu val lys tyr leu lys ser leu his ala met glu

482 512  
GGC CAC GTG ATG ATC GCC TTC TTG CCC ACT ATC CTA AAC CAG CTG TTC CGA GTC CTC ACC  
gly his val met ile ala phe leu pro thr ile leu asn gln leu phe arg val leu thr

542 572  
AGA GCC ACA CAG GAA GAA GTC GCG GTT AAC GTG ACT CGG GTC ATT ATT CAT GTG GTT GCC  
arg ala thr gln glu glu val ala val asn val thr arg val ile ile his val val ala

602 632  
CAG TGC CAT GAG GAA GGA TTG GAG AGC CAC TTG AGG TCA TAT GTT AAG TAC GCG TAT AAG  
gln cys his glu glu gly leu glu ser his leu arg ser tyr val lys tyr ala tyr lys

662 692  
GCT GAG CCA TAT GTT GCC TCT GAA TAC AAG ACA GTG CAT GAA GAA CTG ACC AAA TCC ATG  
ala glu pro tyr val ala ser glu tyr lys thr val his glu glu leu thr lys ser met

FIG. 1

722 ACC ACG ATT CTC AAG CCT TCT GCC GAT TTC CTC ACC AGC AAC AAA CTA CTG AGG TAC TCA  
thr thr ile leu lys pro ser ala asp phe leu thr ser asn lys leu leu arg tyr ser

782 TGG TTT TTC TTT GAT GTA CTG ATC AAA TCT ATG GCT CAG CAT TTG ATA GAG AAC TCC AAA  
trp phe phe phe asp val leu ile lys ser met ala gln his leu ile glu asn ser lys

842 |Cadherin Cleavage| 872 GTT AAG TTG CTG CGA AAC CAG AGA TTT CCT GCA TCC TAT CAT CAT GCA GCG GAA ACC GTT  
val lys leu leu arg asn gln arg phe pro ala ser tyr his his ala ala glu thr val

902 GTA AAT ATG CTG ATG CCA CAC ATC ACT CAG AAG TTT GGA GAT AAT CCA GAG GCA TCT AAG  
val asn met leu met pro his ile thr gln lys phe gly asp asn pro glu ala ser lys

962 AAC GCG AAT CAT AGC CTT GCT GTC TTC ATC AAG AGA TGT TTC ACC TTC ATG GAC AGG GGC  
asn ala asn his ser leu ala val phe ile lys arg cys phe thr phe met asp arg gly

1022 TTT GTC TTC AAG CAG ATC AAC AAC TAC ATT AGC TGT TTT GCT CCT GGA GAC CCA AAG ACC  
phe val phe lys gln ile asn asn tyr ile ser cys phe ala pro gly asp pro lys thr

1082 CTC TTT GAA TAC AAG TTT GAA TTT CTC CGT GTA GTG TGC AAC CAT GAA CAT TAT ATT CCG  
leu phe glu tyr lys phe glu phe leu arg val val cys asn his glu his tyr ile pro

1142 TTG AAC TTA CCA ATG CCA TTT GGA AAA GGC AGG ATT CAA AGA TAC CAA GAC CTC CAG CTT  
leu asn leu pro met pro phe gly lys gly arg ile gln arg tyr gln asp leu gln leu

1202 GAC TAC TCA TTA ACA GAT GAG TTC TGC AGA AAC CAC TTC TTG GTG GGA CTG TTA CTG AGG  
asp tyr ser leu thr asp glu phe cys arg asn his phe leu val gly leu leu leu arg

xxx| 1292 GAG GTG GGG ACA GCC CTC CAG GAG TTC CGG GAG GTC CGT CTG ATC GCC ATC AGT GTG CTC  
glu val gly thr ala leu gln glu phe arg glu val arg leu ile ala ile ser val leu

1322 AAG AAC CTG CTG ATA AAG CAT TCT TTT GAT GAC AGA TAT GCT TCA AGG AGC CAT CAG GCA  
lys asn leu leu ile lys his ser phe asp asp arg tyr ala ser arg ser his gln ala

1382 AGG ATA GCC ACC CTC TAC CTG CCT CTG TTT GGT CTG CTG ATT GAA AAC GTC CAG CGG ATC  
arg ile ala thr leu tyr leu pro leu phe gly leu leu ile glu asn val gln arg ile

1442 AAT GTG AGG GAT GTG TCA CCC TTC CCT GTG AAC GCG GGC ATG ACC GTG AAG GAT GAA TCC  
asn val arg asp val ser pro phe pro val asn ala gly met thr val lys asp glu ser

1472

FIG. 1 (cont.)

1502	1532
CTG GCT CTA CCA GCT GTG AAT CCG CTG GTG	ACG CCG CAG AAG GGA AGC ACC CTG GAC AAC
leu ala leu pro ala val asn pro leu val	thr pro gln lys gly ser thr leu asp asn
1562	1592
AGC CTG CAC AAG GAC CTG CTG GGC GCC ATC	TCC GGC ATT GCT TCT CCA TAT ACA ACC TCA
ser leu his lys asp leu leu gly ala ile	ser gly ile ala ser pro tyr thr thr ser
1622	1652
ACT CCA AAC ATC AAC AGT GTG AGA AAT GCT	GAT TCG AGA GGA TCT CTC ATA AGC ACA GAT
thr pro asn ile asn ser val arg asn ala	asp ser arg gly ser leu ile ser thr asp
1682	1712
TCG GGT AAC AGC CTT CCA GAA AGG AAT AGT	GAG AAG AGC AAT TCC CTG GAT AAG CAC CAA
ser gly asn ser leu pro glu arg asn ser	glu lys ser asn ser leu asp lys his gln
1742	1772
CAA AGT AGC ACA TTG GGA AAT TCC GTG GTT	CGC TGT GAT AAA CTT GAC CAG TCT GAG ATT
gln ser ser thr leu gly asn ser val val	arg cys asp lys leu asp gln ser glu ile
1802	1832
AAG AGC CTA CTG ATG TGT TTC CTC TAC ATC	TTA AAG AGC ATG TCT GAT GAT GCT TTG TTT
lys ser leu leu met cys phe leu tyr ile	leu lys ser met ser asp asp ala leu phe
1862	1892
ACA TAT TGG AAC AAG GCT TCA ACA TCT GAA	CTT ATG GAT TTT TTT ACA ATA TCT GAA GTC
thr tyr trp asn lys ala ser thr ser glu	leu met asp phe phe thr ile ser glu val
1922	1952
TGC CTG CAC CAG TTC CAG TAC ATG GGG AAG	CGA TAC ATA GCC AGG AAC CAG GAG GGG TTG
cys leu his gln phe gln tyr met gly lys	arg tyr ile ala arg asn gln glu gly leu
1982	2012
GGA CCC ATA GTT CAT GAT CGA AAG TCT CAG	ACA TTG CCT GTT TCC CGT AAC AGA ACA GGA
gly pro ile val his asp arg lys ser gln	thr leu pro val ser arg asn arg thr gly
2042	2072
ATG ATG CAT GCC AGA TTG CAG CAG CTG GGC	AGC CTG GAT AAC TCT CTC ACT TTT AAC CAC
met met his ala arg leu gln gln leu gly	ser leu asp asn ser leu thr phe asn his
2102	2132
AGC TAT GGC CAC TCG GAC GCA GAT GTT CTG	CAC CAG TCA TTA CTT GAA GCC AAC ATT GCT
ser tyr gly his ser asp ala asp val leu	his gln ser leu leu glu ala asn ile ala
2162	2192
ACT GAG GTT TGC CTG ACA GCT CTG GAC ACG	CTT TCT CTA TTT ACA TTG GCG TTT AAG AAC
thr glu val cys leu thr ala leu asp thr	leu ser leu phe thr leu ala phe lys asn
2222	2252
CAG CTC CTG GCC GAC CAT GGA CAT AAT CCT	CTC ATG AAA AAA GTT TTT GAT GTC TAC CTG
gln leu leu ala asp his gly his asn pro	leu met lys lys val phe asp val tyr leu
2282	2312
TGT TTT CTT CAA AAA CAT CAG TCT GAA ACG	GCT TTA AAA AAT GTC TTC ACT GCC TTA AGG

FIG. 1 (cont.)





cys ala asp gly leu trp lys ala glu arg tyr glu leu ile ala asp ile tyr lys leu

3182 3212  
ATT TAC AAG GAA CCC AAA CTC ACA CCG CTG TCG GAA ATT TCT CAG AGA CTC CTT AAA CTG  
ile tyr lys glu pro lys leu thr pro leu ser glu ile ser gln arg leu leu lys leu

3302 3332  
CCT AAG GAT CTG GAT TCT AAG TAT GCA TAC ATC CAG GTG ACT CAC GTC ATC CCC TTC TTT  
pro lys asp leu asp ser lys tyr ala tyr ile gln val thr his val ile pro phe phe

3422 3452  
 TTC ATG TTT GAG ATG CCA TTT ACG CAG ACC GGG AAG AGG CAG GGC GGG GTG GAA GAG CAG  
 phe met phe glu met pro phe thr gln thr gly lys arg gln gly gly val glu glu gln

3542 3572 |xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx  
CCT GTC ATG TAC CAG CAC CAC ACT GAC CTG AAC CCC ATC GAG GTG GCC ATT GAC GAG ATG  
pro val met tyr gln his his thr asp leu asn pro ile glu val ala ile asp glu met

3662 xxxxxxxxxxxxxxxxxxxxxxxxxx| 3692  
CTG CAG CTC AAA CTC CAG GGC AGC GTG AGT GTT CAG GTC AAT GCT GGC CCA CTA GCA TAT  
leu gln leu lys leu gln gly ser val ser val gln val asn ala gly pro leu ala tyr

3782		3812	xxxxxxxxxxxxxxxxxxxxxx
CTT AAG GAA GTT TTC AGG CAA TTT GTG GAA GCT TGC GGT CAA GCC		TTA GCG GTA AAC GAA	
leu lys glu val phe arg gln phe val glu ala cys gly gln ala		leu ala val asn glu	

FIG. 1 (cont.)

arg leu ile lys glu asp gln leu glu tyr gln glu glu met lys ala asn tyr arg glu

3902 xxx Coiled-Coil 2 xxxxxxxxxxxxxxxx 3932 xxxxx|

ATG GCG AAG GAG CTT TCT GAA ATC ATG CAT GAG CAG ATC TGC CCC CTG GAG GAG AAG ACG  
met ala lys glu leu ser glu ile met his glu gln ile cys pro leu glu glu lys thr

3962

3992

AGC GTC TTA CCG AAT TCC CTT CAC ATC TTC AAC GCC ATC AGT GGG ACT CCA ACA AGC ACA  
ser val leu pro asn ser leu his ile phe asn ala ile ser gly thr pro thr ser thr

4022

|xxxx PBM xxxx|

ATG GTT CAC GGG ATG ACC AGC TCG TCT TCG GTC GTG TGA TTA CAT CTC ATG GCC CGT GTG  
met val his gly met thr ser ser ser ser val val STP

4082

4112

TGG GGA CTT GCT TTG TCA TTT GCA AAC TCA GGA TGC TTT CCA AAG CCA ATC ACT GGG GAG

4142

4172

ACC GAG CAC AGG GAG GAC CAA GGG GAA GGG GAG AGA AAG GAA ATA AAG AAC AAC GTT ATT

4202

4232

TCT TAA CAG ACT TTC TAT AGG AGT TGT AAG AAG GTG CAC ATA TTT TTT TAA ATC TCA CTG

4262

4292

GCA ATA TTC AAA GTT TTC ATT GTG TCT TAA CAA AGG TGT GGT AGA CAC TCT TGA GCT GGA

4322

4352

CTT AGA TTT TAT TCT TCC TTG CAG AGT AGT GTT AGA ATA GAT GGC CTA CAG AAA AAA AAG

4382

4412

GTT CTG GGA TCT ACA TGG CAG GGA GGG CTG CAC TGA CAT TGA TGC CTG GGG GAC CTT TTG

4442

4472

CCT CGA CTC GTG CCG GAA ATC TGA TCG TAA TCA GGG TAC AGA ACT TAC TAG TTT TGT CTA

4502

4532

GGA GTA TGT TGT ATG ACT AGG ATT TGT GCT ATT ATC TCA TTC AAC AAC ATA GAG CAA GAA

4562

4592

TAG TGA GCT AAC TGA GCT AGA CAC TCA ATT AAT CCG CTA CTG GCT TCA AGT CAG AAC TTT

4622

4652

GTC ATT AAT CAT CGA CTC CGG GAC GGT CAT ATA TGT ATT ACA TTT CTA CAT TTT TAA TAC

4682

4712

TCA CAT GGG CTT ATG CAT TAA GTT TAA TTG TGA TAA ATT TGT GCT GGT CCA GTA TAT GCA

4742

4772

ATA CAC TTT AAT GGT TTA TTC TTG TCA TAA AAA TGT GCA ATA TGG AGA TGT ATA CAA GTC

4802

TTT ACT

FIG. 1 (cont.)

(Nucleotide position for insertions and deletion are found above the Human (h) CLASP-2A line diagram. Numbers are referenced based on hCLASP-2A nucleotide sequence from Figure 1.)

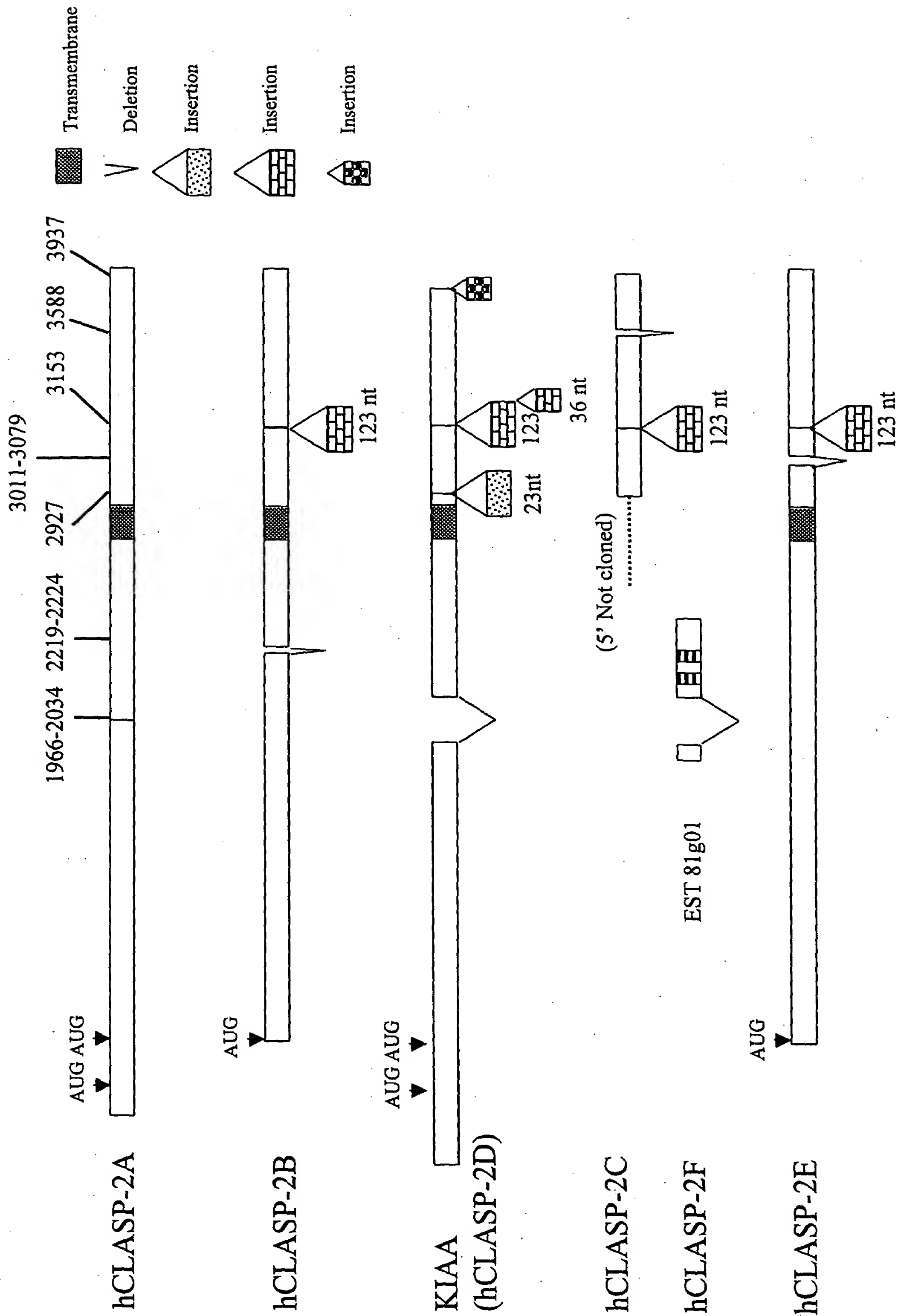


FIG. 2A

2 32  
GTT TTA CAC CAT CAC CAA AAC CCA GAA TTT TAT GAT GAG ATT AAA ATA GAG TTG CCC ACT  
val leu his his his gln asn pro glu phe tyr asp glu ile lys ile glu leu pro thr

62 92  
CAG CTG CAT GAA AAG CAC CAC CTG TTG CTC ACA TTC TTC CAT GTC AGC TGT GAC AAC TCA  
gln leu his glu lys his his leu leu leu thr phe phe his val ser cys asp asn ser

122 152  
AGT AAA GGA AGC ACG AAG AAG AGG GAT GTC GTT GAA ACC CAA GTT GGC TAC TCC TGG CTT  
ser lys gly ser thr lys lys arg asp val val glu thr gln val gly tyr ser trp leu

182 212  
CCC CTC CTG AAA GAC GGA AGG GTG GTG ACA AGC GAG CAG CAC ATC CCG GTC TCG GCG AAC  
pro leu leu lys asp gly arg val val thr ser glu gln his ile pro val ser ala asn

242 272  
CTT CCT TCG GGC TAT CTT GGC TAC CAA GAG CTT GGG ATG GGC AGG CAT TAT GGT CCG GAA  
leu pro ser gly tyr leu gly tyr gln glu leu gly met gly arg his tyr gly pro glu

302 332  
ATT AAA TGG GTA GAT GGA GGC AAG CCA CTG CTG AAA ATT TCC ACT CAT CTG GTT TCT ACA  
ile lys trp val asp gly gly lys pro leu leu lys ile ser thr his leu val ser thr

362 392  
GTG TAT ACT CAG GAT CAG CAT TTA CAT AAT TTT TTC CAG TAC TGT CAG AAA ACC GAA TCT  
val tyr thr gln asp gln his leu his asn phe phe gln tyr cys gln lys thr glu ser

422 452  
GGA GCC CAA GCC TTA GGA AAC GAA CTT GTA AAG TAC CTT AAG AGT CTG CAT GCG ATG GAA  
gly ala gln ala leu gly asn glu leu val lys tyr leu lys ser leu his ala met glu

482 512  
GGC CAC GTG ATG ATC GCC TTC TTG CCC ACT ATC CTA AAC CAG CTG TTC CGA GTC CTC ACC  
gly his val met ile ala phe leu pro thr ile leu asn gln leu phe arg val leu thr

542 572  
AGA GCC ACA CAG GAA GAA GTC GCG GTT AAC GTG ACT CGG GTC ATT ATT CAT GTG GTT GCC  
arg ala thr gln glu glu val ala val asn val thr arg val ile ile his val val ala

602 632  
CAG TGC CAT GAG GAA GGA TTG GAG AGC CAC TTG AGG TCA TAT GTT AAG TAC GCG TAT AAG  
gln cys his glu glu gly leu glu ser his leu arg ser tyr val lys tyr ala tyr lys

662 692  
GCT GAG CCA TAT GTT GCC TCT GAA TAC AAG ACA GTG CAT GAA GAA CTG ACC AAA TCC ATG  
ala glu pro tyr val ala ser glu tyr lys thr val his glu glu leu thr lys ser met

FIG. 2B

722	ACC ACG ATT CTC AAG CCT TCT GCC GAT TTC	752	CTC ACC AGC AAC AAA CTA CTG AGG TAC TCA
	thr thr ile leu lys pro ser ala asp phe		leu thr ser asn lys leu leu arg tyr ser
782	TGG TTT TTC TTT GAT GTA CTG ATC AAA TCT	812	ATG GCT CAG CAT TTG ATA GAG AAC TCC AAA
	trp phe phe phe asp val leu ile lys ser		met ala gln his leu ile glu asn ser lys
842	Cadherin Cleavage		872
	GTT AAG TTG CTG CGA AAC CAG AGA TTT CCT		GCA TCC TAT CAT CAT GCA GCG GAA ACC GTT
	val lys leu leu arg asn gln arg phe pro		ala ser tyr his his ala ala glu thr val
902	GTA AAT ATG CTG ATG CCA CAC ATC ACT CAG	932	AAG TTT GGA GAT AAT CCA GAG GCA TCT AAG
	val asn met leu met pro his ile thr gln		lys phe gly asp asn pro glu ala ser lys
962	AAC GCG AAT CAT AGC CTT GCT GTC TTC ATC	992	AAG AGA TGT TTC ACC TTC ATG GAC AGG GGC
	asn ala asn his ser leu ala val phe ile		lys arg cys phe thr phe met asp arg gly
1022	TTT GTC TTC AAG CAG ATC AAC AAC TAC ATT	1052	AGC TGT TTT GCT CCT GGA GAC CCA AAG ACC
	phe val phe lys gln ile asn asn tyr ile		ser cys phe ala pro gly asp pro lys thr
1082	CTC TTT GAA TAC AAG TTT GAA TTT CTC CGT	1112	GTA GTG TGC AAC CAT GAA CAT TAT ATT CCG
	leu phe glu tyr lys phe glu phe leu arg		val val cys asn his glu his tyr ile pro
1142	TTG AAC TTA CCA ATG CCA TTT GGA AAA GGC	1172	AGG ATT CAA AGA TAC CAA GAC CTC CAG CTT
	leu asn leu pro met pro phe gly lys gly		arg ile gln arg tyr gln asp leu gln leu
1202	GAC TAC TCA TTA ACA GAT GAG TTC TGC AGA	1232	AAC CAC TTC TTG GTG GGA CTG TTA CTG AGG
	asp tyr ser leu thr asp glu phe cys arg		asn his phe leu val gly leu leu leu arg
xxx1	GAG GTG GGG ACA GCC CTC CAG GAG TTC CGG	1292	GAG GTC CGT CTG ATC GCC ATC AGT GTG CTC
	glu val gly thr ala leu gln glu phe arg		glu val arg leu ile ala ile ser val leu
1322	AAG AAC CTG CTG ATA AAG CAT TCT TTT GAT	1352	GAC AGA TAT GCT TCA AGG AGC CAT CAG GCA
	lys asn leu leu ile lys his ser phe asp		asp arg tyr ala ser arg ser his gln ala
1382	AGG ATA GCC ACC CTC TAC CTG CCT CTG TTT	1412	GGT CTG CTG ATT GAA AAC GTC CAG CGG ATC
	arg ile ala thr leu tyr leu pro leu phe		gly leu leu ile glu asn val gln arg ile
1442	AAT GTG AGG GAT GTG TCA CCC TTC CCT GTG	1472	AAC GCG GGC ATG ACC GTG AAG GAT GAA TCC
	asn val arg asp val ser pro phe pro val		asn ala gly met thr val lys asp glu ser

FIG. 2B (cont.)

1502 1532  
CTG GCT CTA CCA GCT GTG AAT CCG CTG GTG ACG CCG CAG AAG GGA AGC ACC CTG GAC AAC  
leu ala leu pro ala val asn pro leu val thr pro gln lys gly ser thr leu asp asn

1562 1592  
AGC CTG CAC AAG GAC CTG CTG GGC GCC ATC TCC GGC ATT GCT TCT CCA TAT ACA ACC TCA  
ser leu his lys asp leu leu gly ala ile ser gly ile ala ser pro tyr thr thr ser

1622 1652  
ACT CCA AAC ATC AAC AGT GTG AGA AAT GCT GAT TCG AGA GGA TCT CTC ATA AGC ACA GAT  
thr pro asn ile asn ser val arg asn ala asp ser arg gly ser leu ile ser thr asp

1682 1712  
TCG GGT AAC AGC CTT CCA GAA AGG AAT AGT GAG AAG AGC AAT TCC CTG GAT AAG CAC CAA  
ser gly asn ser leu pro glu arg asn ser glu lys ser asn ser leu asp lys his gln

1742 1772  
CAA AGT AGC ACA TTG GGA AAT TCC GTG GTT CGC TGT GAT AAA CTT GAC CAG TCT GAG ATT  
gln ser ser thr leu gly asn ser val val arg cys asp lys leu asp gln ser glu ile

1802 1832  
AAG AGC CTA CTG ATG TGT TTC CTC TAC ATC TTA AAG AGC ATG TCT GAT GAT GCT TTG TTT  
lys ser leu leu met cys phe leu tyr ile leu lys ser met ser asp asp ala leu phe

1862 1892  
ACA TAT TGG AAC AAG GCT TCA ACA TCT GAA CTT ATG GAT TTT TTT ACA ATA TCT GAA GTC  
thr tyr trp asn lys ala ser thr ser glu leu met asp phe phe thr ile ser glu val

1922 1952  
TGC CTG CAC CAG TTC CAG TAC ATG GGG AAG CGA TAC ATA GCC AGG AAC CAG GAG GGG TTG  
cys leu his gln phe gln tyr met gly lys arg tyr ile ala arg asn gln glu gly leu

1982 xxxxxxxx deleted in CLASP-2D(KIAA1058) xxxxxxxxxxxxxxxxxxxxxxxx!  
GGA CCC ATA GTT CAT GAT CGA AAG TCT CAG ACA TTG CCT GTT TCC CGT AAC AGA ACA GGA  
gly pro ile val his asp arg lys ser gln thr leu pro val ser arg asn arg thr gly

2042 2072  
ATG ATG CAT GCC AGA TTG CAG CAG CTG GGC AGC CTG GAT AAC TCT CTC ACT TTT AAC CAC  
met met his ala arg leu gln gln leu gly ser leu asp asn ser leu thr phe asn his

2102 2132  
AGC TAT GGC CAC TCG GAC GCA GAT GTT CTG CAC CAG TCA TTA CTT GAA GCC AAC ATT GCT  
ser tyr gly his ser asp ala asp val leu his gln ser leu leu glu ala asn ile ala

2162 2192  
ACT GAG GTT TGC CTG ACA GCT CTG GAC ACG CTT TCT CTA TTT ACA TTG GCG TTT AAG AAC  
thr glu val cys leu thr ala leu asp thr leu ser leu phe thr leu ala phe lys asn

in HC2B  
xxxl 2252  
CAG CTC CTG GCC GAC CAT GGA CAT AAT CCT CTC ATG AAA AAA GTT TTT GAT GTC TAC CTG  
gln leu leu ala asp his gly his asn pro leu met lys lys val phe asp val tyr leu

Deleted  
Ixxx



2282	2312	
TGT TTT CTT CAA AAA CAT CAG TCT GAA ACG	GCT TTA AAA AAT GTC TTC ACT GCC TTA AGG	
cys phe leu gln lys his gln ser glu thr	ala leu lys asn val phe thr ala leu arg	
2342	2372	
TCC TTA ATT TAT AAG TTT CCC TCA ACA TTC	TAT GAA GGG AGA GCG GAC ATG TGT GCG GCT	
ser leu ile tyr lys phe pro ser thr phe	tyr glu gly arg ala asp met cys ala ala	
2402	2432	
CTG TGT TAC GAG ATT CTC AAG TGC TGT AAC	TCC AAG CTG AGC TCC ATC AGG ACG GAG GCC	
leu cys tyr glu ile leu lys cys cys asn	ser lys leu ser ser ile arg thr glu ala	
2462	2492	
TCC CAG CTG CTC TAC TTC CTG ATG AGG AAC	AAC TTT GAT TAC ACT GGA AAG AAG TCC TTT	
ser gln leu leu tyr phe leu met arg asn	asn phe asp tyr thr gly lys lys ser phe	
2522	2552	
GTC CGG ACA CAT TTG CAA GTC ATC ATA TCT	GTC AGC CAG CTG ATA GCA GAC GTT GTT GGC	
val arg thr his leu gln val ile ile ser	val ser gln leu ile ala asp val val gly	
2582	2612	
ATT GGG GAA ACC AGA TTC CAG CAG TCC CTG	TCC ATC ATC AAC AAC TGT GCC AAC AGT GAC	
ile gly glu thr arg phe gln gln ser leu	ser ile ile asn asn cys ala asn ser asp	
2642	2672	
GGG CTT ATT AAG CAC ACC AGC TTC TCC TCT	GAT GTG AAG GAC TTA ACC AAA AGG ATA CGC	
arg leu ile lys his thr ser phe ser ser	asp val lys asp leu thr lys arg ile arg	
2702	2732	
ACG GTG CTA ATG GCC ACC GCC CAG ATG AAG	GAG CAT GAG AAC GAC CCA GAG ATG CTG GTG	
thr val leu met ala thr ala gln met lys	glu his glu asn asp pro glu met leu val	
2762	2792	
GAC CTC CAG TAC AGC CTG GCC AAA TCC TAT	GCC AGC ACG CCC GAG CTC AGG AAG ACG TGG	
asp leu gln tyr ser leu ala lys ser tyr	ala ser thr pro glu leu arg lys thr trp	
2822	2852	lxxxxxxxxxxxxxxxx Predicted
CTC GAC AGC ATG GCC AGG ATC CAT GTC AAA AAT	GGC GAT CTC TCA GAG GCA GCA ATG TGC	
leu asp ser met ala arg ile his val lys	asn gly asp leu ser glu ala ala met cys	

[Additional and differential exon usage found at position 2927 consisting of 69 nucleotides. This entire sequence is found in Human CLASP-2D (KIAA1058) and not other isoforms of CLASP-2. It has a sequence of:  
AAGCAGTCCAGTGGGAGCCGCCCTTCTCCCCACAGCCATAGCGCCTGCCTGAGGAGGAGCCGGGGAG]

Transmembrane Domain xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx|

TAT GTC CAC GTA ACA GCC CTA GTG GCA GAA TAT CTC ACA CGG AAA GGC GTG TTT AGA CAA  
tyr val his val thr ala leu val ala glu tyr leu thr arg lys gly val phe arg gln

2942	2972
GGA TGC ACC GCC TTC AGG GTC ATT ACC CCA AAC	ATC GAC GAG GAG GCC TCC ATG ATG GAA
gly cys thr ala phe arg val ile thr pro	asn ile asp glu glu ala ser met met glu

3002 |xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx Sequence deleted in CLASP-2E xxxxx  
GAC GTG GGG ATG CAG GAT GTC CAT TTC AAC GAG GAT GTG CTG ATG GAG CTC CTT GAG CAG  
asp val gly met gln asp val his phe asn glu asp val leu met glu leu leu glu gln

3062 xxxxxxxxxxxxxxxxxxxxxxxx| 3092  
TGC GCA GAT GGA CTC TGG AAA GCC GAG CGC TAC GAG CTC ATC GCC GAC ATC TAC AAA CTT  
cys ala asp gly leu trp lys ala glu arg tyr glu leu ile ala asp ile tyr lys leu

[[Additional and differential exon usage found at position 3153. The  
entire sequence below is found in Human CLASP-2D. Underlined sequence is  
found in Human CLASP-2B, 2C and 2E.

TGAGAGGCTGGCCCATCTGTATGACACGCTGCACCGGGCCTACAGCAAAGTGACCGAGGTCAT  
GCACTCGGGCCGAGGCTTCTGGGGACCTACTTCCGGGTAGCCTTCTTCGGGCAGGCAGCGCAATACCAGTTT  
ACAGACAGTGAAACAGATGTGGAGGGATT]



3122 3155  
ATC ATC CCC ATT TAT GAG AAG CGG AGG GAT TTC TTT GAA GAT GAA GAT GGA AAG GAG TAT  
ile ile pro ile tyr glu lys arg arg asp phe phe glu asp glu asp gly lys glu tyr

3182 3212  
ATT TAC AAG GAA CCC AAA CTC ACA CCG CTG TCG GAA ATT TCT CAG AGA CTC CTT AAA CTG  
ile tyr lys glu pro lys leu thr pro leu ser glu ile ser gln arg leu leu lys leu

3242 3272  
TAC TCG GAT AAA TTT GGT TCT GAA AAT GTC AAA ATG ATA CAG GAT TCT GGC AAG GTC AAC  
tyr ser asp lys phe gly ser glu asn val lys met ile gln asp ser gly lys val asn

3302 3332  
CCT AAG GAT CTG GAT TCT AAG TAT GCA TAC ATC CAG GTG ACT CAC GTC ATC CCC TTC TTT  
pro lys asp leu asp ser lys tyr ala tyr ile gln val thr his val ile pro phe phe

3362 3392  
GAC GAA AAA GAG TTG CAA GAA AGG AAA ACA GAG TTT GAG AGA TCC CAC AAC ATC CGC CGC  
asp glu lys glu leu gln glu arg lys thr glu phe glu arg ser his asn ile arg arg

3422 3452  
TTC ATG TTT GAG ATG CCA TTT ACG CAG ACC GGG AAG AGG CAG GGC GGG GTG GAA GAG CAG  
phe met phe glu met pro phe thr gln thr gly lys arg gln gly gly val glu glu gln

3482 3512  
TGC AAA CGG CGC ACC ATC CTG ACA GCC ATA CAC TGC TTC CCT TAT GTG AAG AAG CGC ATC  
cys lys arg arg thr ile leu thr ala ile his cys phe pro tyr val lys lys arg ile

Two nucleotide deletion (nts 3586 and 3587) found in Human CLASP-2C



3542 3572 |xxx|  
CCT GTC ATG TAC CAG CAC CAC ACT GAC CTG AAC CCC ATC GAG GTG GCC ATT GAC GAG ATG  
pro val met tyr gln his his thr asp leu asn pro ile glu val ala ile asp glu met

Fig. 2B(cont.)

3602 3632  
AGT AAG AAG GTG GCG GAG CTC CGG CAG CTG TGC TCC TCG GCC GAG GTG GAC ATG ATC AAA  
ser lys lys val ala glu leu arg gln leu cys ser ser ala glu val asp met ile lys

3662 3692  
CTG CAG CTC AAA CTC CAG GGC AGC GTG AGT GTT CAG GTC AAT GCT GGC CCA CTA GCA TAT  
leu gln leu lys leu gln gly ser val ser val gln val asn ala gly pro leu ala tyr

3722 3752  
GCG CGA GCT TTC TTA GAT GAT ACA AAC ACA AAG CGA TAT CCT GAC AAT AAA GTG AAG CTG  
ala arg ala phe leu asp asp thr asn thr lys arg tyr pro asp asn lys val lys leu

3782 3812  
CTT AAG GAA GTT TTC AGG CAA TTT GTG GAA GCT TGC GGT CAA GCC TTA GCG GTA AAC GAA  
leu lys glu val phe arg gln phe val glu ala cys gly gln ala leu ala val asn glu

3842 3872  
CGT CTG ATT AAA GAA GAC CAG CTC GAG TAT CAG GAA GAA ATG AAA GCC AAC TAC AGG GAA  
arg leu ile lys glu asp gln leu glu tyr gln glu glu met lys ala asn tyr arg glu

Insertion of 8 nucleotides found only in Human CLASP-2D with sequence: CTGGGATG

3902 3932  
ATG GCG AAG GAG CTT TCT GAA ATC ATG CAT GAG CAG ATC TGC CCC CTG GAG GAG AAG ACG  
met ala lys glu leu ser glu ile met his glu gln ile cys pro leu glu glu lys thr

3962 3992  
AGC GTC TTA CCG AAT TCC CTT CAC ATC TTC AAC GCC ATC AGT GGG ACT CCA ACA AGC ACA  
ser val leu pro asn ser leu his ile phe asn ala ile ser gly thr pro thr ser thr

4022 |xxxxx PBM xxxxx|  
ATG GTT CAC GGG ATG ACC AGC TCG TCT TCG GTC GTG TGA TTA CAT CTC ATG GCC CGT GTG  
met val his gly met thr ser ser ser ser val val STP

4082 4112  
TGG GGA CTT GCT TTG TCA TTT GCA AAC TCA GGA TGC TTT CCA AAG CCA ATC ACT GGG GAG

4142 4172  
ACC GAG CAC AGG GAG GAC CAA GGG GAA GGG GAG AGA AAG GAA ATA AAG AAC AAC GTT ATT

4202 4232  
TCT TAA CAG ACT TTC TAT AGG AGT TGT AAG AAG GTG CAC ATA TTT TTT TAA ATC TCA CTG

4262 4292  
GCA ATA TTC AAA GTT TTC ATT GTG TCT TAA CAA AGG TGT GGT AGA CAC TCT TGA GCT GGA

4322 4352  
CTT AGA TTT TAT TCT TCC TTG CAG AGT AGT GTT AGA ATA GAT GGC CTA CAG AAA AAA AAG

4382 4412  
GTT CTG GGA TCT ACA TGG CAG GGA GGG CTG CAC TGA CAT TGA TGC CTG GGG GAC CTT TTG

4442 4472

Fig. 2B (cont.)

CCT CGA CTC GTG CCG GAA ATC TGA TCG TAA TCA GGG TAC AGA ACT TAC TAG TTT TGT CTA

4502

4532

GGA GTA TGT TGT ATG ACT AGG ATT TGT GCT ATT ATC TCA TTC AAC AAC ATA GAG CAA GAA

4562

4592

TAG TGA GCT AAC TGA GCT AGA CAC TCA ATT AAT CCG CTA CTG GCT TCA AGT CAG AAC TTT

4622

4652

GTC ATT AAT CAT CGA CTC CGG GAC GGT CAT ATA TGT ATT ACA TTT CTA CAT TTT TAA TAC

4682

4712

TCA CAT GGG CTT ATG CAT TAA GTT TAA TTG TGA TAA ATT TGT GCT GGT CCA GTA TAT GCA

4742

4772

ATA CAC TTT AAT GGT TTA TTC TTG TCA TAA AAA TGT GCA ATA TGG AGA TGT ATA CAA GTC

4802

TTT ACT

09687837 101300

Fig. 2B(cont.)

HC2A -----  
HC2-80 -----  
HC2B -----  
HC2C -----  
HC2D-KIAA1058 GCATCTGGAAATCTTGACAAAAATGCCAGATTTTCTGCCATCTACAGGCAAGACAGCAAT  
HC2E -----  
HC2F -----

HC2A -----  
HC2-80 -----  
HC2B -----  
HC2C -----  
HC2D-KIAA1058 AAGCTATCCAATGATGACATGCTCAAGTTACTTGCAGACTTTCGGAAACCTGAGAAGATG  
HC2E -----  
HC2F -----

HC2A -----  
HC2-80 -----  
HC2B -----  
HC2C -----  
HC2D-KIAA1058 GCTAAGCTCCCAGTGATTTTAGGCAATCTAGACATTACAATTGATAATGTTTCCTCAGAC  
HC2E -----  
HC2F -----

HC2A -----  
HC2-80 -----  
HC2B -----  
HC2C -----  
HC2D-KIAA1058 TTCCCTAATTATGTTAATTCATCATACTCCACAAAACAATTTGAAACCTGCAGTAAA  
HC2E -----  
HC2F -----

HC2A -----  
HC2-80 -----  
HC2B -----  
HC2C -----  
HC2D-KIAA1058 ACTCCCATCACGTTTGAAGTGGAGGAATTTGTGCCCTGCATACCAAACACACTCAGCCT  
HC2E -----  
HC2F -----

FIG. 3A

HC2A	-----
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	TACACCATCTACACCAATCACCTTTACGTTTATCCTAAGTACTTGAAATACGACAGTCAG
HC2E	-----
HC2F	-----

HC2A	-----
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	AAGTCTTTTGCCAAGGCTAGAAATATTGCGATTTGCATTGAATTCAAAGATTCAGATGAG
HC2E	-----
HC2F	-----

HC2A	-----
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	GAAGACTCTCAGCCCCCTTAAGTGCATTTATGGCAGACCTGGTGGGCCAGTTTTTCACAAGA
HC2E	-----
HC2F	-----

HC2A	-----AGTTTTACACCATCACCAAACCCAGAATTTTATGATGAGATTAAA
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	AGCGCCTTTGCTGCAGTTTTACACCATCACCAAACCCAGAATTTTATGATGAGATTAAA
HC2E	-----
HC2F	-----

HC2A	ATAGAGTTGCCCACTCAGCTGCATGAAAAGCACCACTGTTGCTCACATTCTTCCATGTC
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	ATAGAGTTGCCCACTCAGCTGCATGAAAAGCACCACTGTTGCTCACATTCTTCCATGTC
HC2E	-----
HC2F	-----

HC2A	AGCTGTGACAACTCAAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTT
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	AGCTGTGACAACTCAAGTAAAGGAAGCACGAAGAAGAGGGATGTCGTTGAAACCCAAGTT
HC2E	-----
HC2F	-----

FIG. 3A (cont.)



HC2A	GGCTACTCCTGGCTTCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATC
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	GGCTACTCCTGGCTTCCCCTCCTGAAAGACGGAAGGGTGGTGACAAGCGAGCAGCACATC
HC2E	-----
HC2F	-----

HC2A	CCGGTCTCGGCGAACCTTCCTTCGGGCTATCTTGGCTACCAAGAGCTTGGGATGGGCAGG
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	CCGGTCTCGGCGAACCTTCCTTCGGGCTATCTTGGCTACCAGGAGCTTGGGATGGGCAGG
HC2E	-----
HC2F	-----

HC2A	CATTATGGTCCGGAAATTAAATGGGTTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACT
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	CATTATGGTCCGGAAATTAAATGGGTTAGATGGAGGCAAGCCACTGCTGAAAATTTCCACT
HC2E	-----
HC2F	-----

HC2A	CATCTGGTTTCTACAGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGT
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	CATCTGGTTTCTACAGTGTATACTCAGGATCAGCATTTACATAATTTTTTCCAGTACTGT
HC2E	-----
HC2F	-----

HC2A	CAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAACGAACTTGTAAGTACCTTAAGAGT
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	CAGAAAACCGAATCTGGAGCCCAAGCCTTAGGAAACGAACTTGTAAGTACCTTAAGAGT
HC2E	-----
HC2F	-----

HC2A	CTGCATGCGATGGAAGGCCACGTGATGATCGCCTTCTTGCCCACTATCCTAAACCAGCTG
HC2-80	-----
HC2B	-----GCGATGGAAGGCCACGTGATGATCGCCTTCTTGCCCACTATCCTAAACCAGCTG
HC2C	-----
HC2D-KIAA1058	CTGCATGCGATGGAAGGCCACGTGATGATCGCCTTCTTGCCCACTATCCTAAACCAGCTG
HC2E	-----GCGATGGAAGGCCACGTGATGATCGCCTTCTTGCCCACTATCCTAAACCAGCTG
HC2F	-----

Fig. 3A (cont.)

HC2A	TTCCGAGTCCTCACCAGAGCCACACAGGAAGAAGTCGCGGTTAACGTGACTCGGGTCATT
HC2-80	-----
HC2B	TTCCGAGTCCTCACCAGAGCCACACAGGAAGAAGTCGCGGTTAACGTGACTCGGGTCATT
HC2C	-----
HC2D-KIAA1058	TTCCGAGTCCTCACCAGAGCCACACAGGAAGAAGTCGCGGTTAACGTGACTCGGGTCATT
HC2E	TTCCGAGTCCTCACCAGAGCCACACAGGAAGAAGTCGCGGTTAACGTGACTCGGGTCATT
HC2F	-----
HC2A	ATTCATGTGGTTGCCCAGTGCCATGAGGAAGGATTGGAGAGCCACTTGAGGTCATATGTT
HC2-80	-----
HC2B	ATTCATGTGGTTGCCCAGTGCCATGAGGAAGGATTGGAGAGCCACTTGAGGTCATATGTT
HC2C	-----
HC2D-KIAA1058	ATTCATGTGGTTGCCCAGTGCCATGAGGAAGGATTGGAGAGCCACTTGAGGTCATATGTT
HC2E	ATTCATGTGGTTGCCCAGTGCCATGAGGAAGGATTGGAGAGCCACTTGAGGTCATATGTT
HC2F	-----
HC2A	AAGTACGCGTATAAGGCTGAGCCATATGTTGCCTCTGAATACAAGACAGTGCATGAAGAA
HC2-80	-----
HC2B	AAGTACGCGTATAAGGCTGAGCCATATGTTGCCTCTGAATACAAGACAGTGCATGAAGAA
HC2C	-----
HC2D-KIAA1058	AAGTACGCGTATAAGGCTGAGCCATATGTTGCCTCTGAATACAAGACAGTGCATGAAGAA
HC2E	AAGTACGCGTATAAGGCTGAGCCATATGTTGCCTCTGAATACAAGACAGTGCATGAAGAA
HC2F	-----
HC2A	CTGACCAAATCCATGACCACGATTCTCAAGCCTTCTGCCGATTTCTCACCAGCAACAAA
HC2-80	-----
HC2B	CTGACCAAATCCATGACCACGATTCTCAAGCCTTCTGCCGATTTCTCACCAGCAACAAA
HC2C	-----
HC2D-KIAA1058	CTGACCAAATCCATGACCACGATTCTCAAGCCTTCTGCCGATTTCTCACCAGCAACAAA
HC2E	CTGACCAAATCCATGACCACGATTCTCAAGCCTTCTGCCGATTTCTCACCAGCAACAAA
HC2F	-----
HC2A	CTACTGAGGTACTCATGGTTTTTCTTTGATGTACTGATCAAATCTATGGCTCAGCATTG
HC2-80	-----
HC2B	CTACTGAGGTACTCATGGTTTTTCTTTGATGTACTGATCAAATCTATGGCTCAGCATTG
HC2C	-----
HC2D-KIAA1058	CTACTGAAGTACTCATGGTTTTTCTTTGATGTACTGATCAAATCTATGGCTCAGCATTG
HC2E	CTACTGAGGTACTCATGGTTTTTCTTTGATGTACTGATCAAATCTATGGCTCAGCATTG
HC2F	-----
HC2A	ATAGAGAACTCCAAAGTTAAGTTGCTGCGAAACCAGAGATTTCTTGCATCCTATCATCAT
HC2-80	-----
HC2B	ATAGAGAACTCCAAAGTTAAGTTGCTGCGAAACCAGAGATTTCTTGCATCCTATCATCAT
HC2C	-----
HC2D-KIAA1058	ATAGAGAACTCCAAAGTTAAGTTGCTGCGAAACCAGAGATTTCTTGCATCCTATCATCAT
HC2E	ATAGAGAACTCCAAAGTTAAGTTGCTGCGAAACCAGAGATTTCTTGCATCCTATCATCAT
HC2F	-----

Fig. 3A (cont.)

HC2A	GCAGCGGAAACCGTTGTAAATATGCTGATGCCACACATCACTCAGAAGTTTGGAGATAAT
HC2-80	-----
HC2B	GCAGCGGAAACCGTTGTAAATATGCTGATGCCACACATCACTCAGAAGTTTGGAGATAAT
HC2C	-----
HC2D-KIAA1058	GCAGTGGAAACCGTTGTAAATATGCTGATGCCACACATCACTCAGAAGTTTGGAGATAAT
HC2E	GCAGCGGAAACCGTTGTAAATATGCTGATGCCACACATCACTCAGAAGTTTGGAGATAAT
HC2F	-----
HC2A	CCAGAGGCATCTAAGAACGCGAATCATAGCCTTGCTGTCTTCATCAAGAGATGTTTCACC
HC2-80	-----
HC2B	CCAGAGGCATCTAAGAACGCGAATCATAGCCTTGCTGTCTTCATCAAGAGATGTTTCACC
HC2C	-----
HC2D-KIAA1058	CCAGAGGCATCTAAGAACGCGAATCATAGCCTTGCTGTCTTCATCAAGAGATGTTTCACC
HC2E	CCAGAGGCATCTAAGAACGCGAATCATAGCCTTGCTGTCTTCATCAAGAGATGTTTCACC
HC2F	-----
HC2A	TTCATGGACAGGGGCTTTGTCTTCAAGCAGATCAACAACACTACATTAGCTGTTTTGCTCCT
HC2-80	-----
HC2B	TTCATGGACAGGGGCTTTGTCTTCAAGCAGATCAACAACACTACATTAGCTGTTTTGCTCCT
HC2C	-----
HC2D-KIAA1058	TTCATGGACAGGGGCTTTGTCTTCAAGCAGATCAACAACACTACATTAGCTGTTTTGCTCCT
HC2E	TTCATGGACAGGGGCTTTGTCTTCAAGCAGATCAACAACACTACATTAGCTGTTTTGCTCCT
HC2F	-----
HC2A	GGAGACCCAAAGACCCTCTTTGAATACAAGTTTGAATTTCTCCGTGTAGTGTGCAACCAT
HC2-80	-----
HC2B	GGAGACCCAAAGACCCTCTTTGAATACAAGTTTGAATTTCTCCGTGTAGTGTGCAACCAT
HC2C	-----
HC2D-KIAA1058	GGAGACCCAAAGACCCTCTTTGAATACAAGTTTGAATTTCTCCGTGTAGTGTGCAACCAT
HC2E	GGAGACCCAAAGACCCTCTTTGAATACAAGTTTGAATTTCTCCGTGTAGTGTGCAACCAT
HC2F	-----
HC2A	GAACATTATATTCCGTTGAACTTACCAATGCCATTTGGAAAAGGCAGGATTCAAAGATAC
HC2-80	-----
HC2B	GAACATTATATTCCGTTGAACTTACCAATGCCATTTGGAAAAGGCAGGATTCAAAGATAC
HC2C	-----
HC2D-KIAA1058	GAACATTATATTCCGTTGAACTTACCAATGCCATTTGGAAAAGGCAGGATTCAAAGATAC
HC2E	GAACATTATATTCCGTTGAACTTACCAATGCCATTTGGAAAAGGCAGGATTCAAAGATAC
HC2F	-----
HC2A	CAAGACCTCCAGCTTGACTACTCATTAACAGATGAGTTCTGCAGAAACCACTTCTTGGTG
HC2-80	-----TCCAGCTTGACTACTCATTAACAGATGAGTTCTGCAGAAACCACTTCTTGGTG
HC2B	CAAGACCTCCAGCTTGACTACTCATTAACAGATGAGTTCTGCAGAAACCACTTCTTGGTG
HC2C	-----
HC2D-KIAA1058	CAAGACCTCCAGCTTGACTACTCATTAACAGATGAGTTCTGCAGAAACCACTTCTTGGTG
HC2E	CAAGACCTCCAGCTTGACTACTCATTAACAGATGAGTTCTGCAGAAACCACTTCTTGGTG
HC2F	-----

Fig. 3A (cont.)

HC2A	GGACTGTTACTGAGGGAGGTGGGGACAGCCCTCCAGGAGTTCCGGGAGGTCCGTCTGATC
HC2-80	GGACTGTTACTGAGGGAGGTGGGGACAGCCCTCCAGGAGTTCCGGGAGGTCCGTCTGATC
HC2B	GGACTGTTACTGAGGGAGGTGGGGACAGCCCTCCAGGAGTTCCGGGAGGTCCGTCTGATC
HC2C	-----
HC2D-KIAA1058	GGACTGTTACTGAGGGAGGTGGGGACAGCCCTCCAGGAGTTCCGGGAGGTCCGTCTGATC
HC2E	GGACTGTTACTGAGGGAGGTGGGGACAGCCCTCCAGGAGTTCCGGGAGGTCCGTCTGATC
HC2F	-----
HC2A	GCCATCAGTGTGCTCAAGAACCTGCTGATAAAGCATTCTTTTGATGACAGATATGCTTCA
HC2-80	GCCATCAGTGTGCTCAAGAACCTGCTGATAAAGCATTCTTTTGATGACAGATATGCTTCA
HC2B	GCCATCAGTGTGCTCAAGAACCTGCTGATAAAGCATTCTTTTGATGACAGATATGCTTCA
HC2C	-----
HC2D-KIAA1058	GCCATCAGTGTGCTCAAGAACCTGCTGATAAAGCATTCTTTTGATGACAGATATGCTTCA
HC2E	GCCATCAGTGTGCTCAAGAACCTGCTGATAAAGCATTCTTTTGATGACAGATATGCTTCA
HC2F	-----
HC2A	AGGAGCCATCAGGCAAGGATAGCCACCCTCTACCTGCCTCTGTTTGGTCTGCTGATTGAA
HC2-80	AGGAGCCATCAGGCAAGGATAGCCACCCTCTACCTGCCTCTGTTTGGTCTGCTGATTGAA
HC2B	AGGAGCCATCAGGCAAGGATAGCCACCCTCTACCTGCCTCTGTTTGGTCTGCTGATTGAA
HC2C	-----
HC2D-KIAA1058	AGGAGCCATCAGGCAAGGATAGCCACCCTCTACCTGCCTCTGTTTGGTCTGCTGATTGAA
HC2E	AGGAGCCATCAGGCAAGGATAGCCACCCTCTACCTGCCTCTGTTTGGTCTGCTGATTGAA
HC2F	-----
HC2A	AACGTCCAGCGGATCAATGTGAGGGATGTGTCACCCTTCCCTGTGAACGCGGGCATGACC
HC2-80	AACGTCCAGCGGATCAATGTGAGGGATGTGTCACCCTTCCCTGTGAACGCGGGCATGACC
HC2B	AACGTCCAGCGGATCAATGTGAGGGATGTGTCACCCTTCCCTGTGAACGCGGGCATGACC
HC2C	-----
HC2D-KIAA1058	AACGTCCAGCGGATCAATGTGAGGGATGTGTCACCCTTCCCTGTGAACGCGGGCATGACT
HC2E	AACGTCCAGCGGATCAATGTGAGGGATGTGTCACCCTTCCCTGTGAACGCGGGCATGACC
HC2F	-----
HC2A	GTGAAGGATGAATCCCTGGCTCTACCAGCTGTGAATCCGCTGGTGACGCCGCAGAAGGGA
HC2-80	GTGAAGGATGAATCCCTGGCTCTACCAGCTGTGAATCCGCTGGTGACGCCGCAGAAGGGA
HC2B	GTGAAGGATGAATCCCTGGCTCTACCAGCTGTGAATCCGCTGGTGACGCCGCAGAAGGGA
HC2C	-----
HC2D-KIAA1058	GTGAAGGATGAATCCCTGGCTCTACCAGCTGTGAATCCGCTGGTGACGCCGCAGAAGGGA
HC2E	GTGAAGGATGAATCCCTGGCTCTACCAGCTGTGAATCCGCTGGTGACGCCGCAGAAGGGA
HC2F	-----
HC2A	AGCACCTGGACAACAGCCTGCACAAGGACCTGCTGGGCGCCATCTCCGGCATTGCTTCT
HC2-80	AGCACCTGGACAACAGCCTGCACAAGGACCTGCTGGGCGCCATCTCCGGCATTGCTTCT
HC2B	AGCACCTGGACAACAGCCTGCACAAGGACCTGCTGGGCGCCATCTCCGGCATTGCTTCT
HC2C	-----
HC2D-KIAA1058	AGCACCTGGACAACAGCCTGCACAAGGACCTGCTGGGCGCCATCTCCGGCATTGCTTCT
HC2E	AGCACCTGGACAACAGCCTGCACAAGGACCTGCTGGGCGCCATCTCCGGCATTGCTTCT
HC2F	-----

Fig. 3A (cont.)

HC2A	CCATATACAACCTCAACTCCAAACATCAACAGTGTGAGAAATGCTGATTTCGAGAGGATCT
HC2-80	CCATATACAACCTCAACTCCAAACATCAACAGTGTGAGAAATGCTGATTTCGAGAGGATCT
HC2B	CCATATACAACCTCAACTCCAAACATCAACAGTGTGAGAAATGCTGATTTCGAGAGGATCT
HC2C	-----
HC2D-KIAA1058	CCATATACAACCTCAACTCCAAACATCAACAGTGTGAGAAATGCTGATTTCGAGAGGATCT
HC2E	CCATATACAACCTCAACTCCAAACATCAACAGTGTGAGAAATGCTGATTTCGAGAGGATCT
HC2F	-----GCTGATTTCGAGAGGATCT
HC2A	CTCATAAGCACAGATTTCGGGTAACAGCCTTCCAGAAAGGAATAGTGAGAAGAGCAATTCC
HC2-80	CTCATAAGCACAGATTTCGGGTAACAGCCTTCCAGAAAGGAATAGTGAGAAGAGCAATTCC
HC2B	CTCATAAGCACAGATTTCGGGTAACAGCCTTCCAGAAAGGAATAGTGAGAAGAGCAATTCC
HC2C	-----
HC2D-KIAA1058	CTCATAAGCACAGATTTCGGGTAACAGCCTTCCAGAAAGGAATAGTGAGAAGAGCAATTCC
HC2E	CTCATAAGCACAGATTTCGGGTAACAGCCTTCCAGAAAGGAATAGTGAGAAGAGCAATTCC
HC2F	CTCATAAGCACAGATTTCGGGTAACAGCCTTCCAGAAAGGAATAGTGAGAAGAGCAATTCC
HC2A	CTGGATAAGCACCAACAAAGTAGCACATTGGGAAATTCCGTGGTTCGCTGTGATAAACTT
HC2-80	CTGGATAAGCACCAACAAAGTAGCACATTGGGAAATTCCGTGGTTCGCTGTGATAAACTT
HC2B	CTGGATAAGCACCAACAAAGTAGCACATTGGGAAATTCCGTGGTTCGCTGTGATAAACTT
HC2C	-----
HC2D-KIAA1058	CTGGATAAGCACCAACAAAGTAGCACATTGGGAAATTCCGTGGTTCGCTGTGATAAACTT
HC2E	CTGGATAAGCACCAACAAAGTAGCACATTGGGAAATTCCGTGGTTCGCTGTGATAAACTT
HC2F	CTGGATAAGCACCAACAAAGTAGCACATTGGGAAATTCCGTGGTTCGCTGTGATAAACTT
HC2A	GACCAGTCTGAGATTAAGAGCCTACTGATGTGTTTCCTCTACATCTTAAAGAGCATGTCT
HC2-80	GACCAGTCTGAGATTAAGAGCCTACTGATGTGTTTCCTCTACATCTTAAAGAGCATGTCT
HC2B	GACCAGTCTGAGATTAAGAGCCTACTGATGTGTTTCCTCTACATCTTAAAGAGCATGTCT
HC2C	-----
HC2D-KIAA1058	GACCAGTCTGAGATTAAGAGCCTACTGATGTGTTTCCTCTACATCTTAAAGAGCATGTCT
HC2E	GACCAGTCTGAGATTAAGAGCCTACTGATGTGTTTCCTCTACATCTTAAAGAGCATGTCT
HC2F	GACCAGTCTGAGATTAAGAGCCTACTGATGTGTTTCCTCTACATCTTAAAGAGCATGTCT
HC2A	GATGATGCTTTGTTTACATATTGGAACAAGGCTTCAACATCTGAACTTATGGATTTTTTTT
HC2-80	GATGATGCTTTGTTTACATATTGGAACAAGGCTTCAACATCTGAACTTATGGATTTTTTTT
HC2B	GATGATGCTTTGTTTACATATTGGAACAAGGCTTCAACATCTGAACTTATGGATTTTTTTT
HC2C	-----
HC2D-KIAA1058	GATGATGCTTTGTTTACATATTGGAACAAGGCTTCAACATCTGAACTTATGGATTTTTTTT
HC2E	GATGATGCTTTGTTTACATATTGGAACAAGGCTTCAACATCTGAACTTATGGATTTTTTTT
HC2F	GATGATGCTTTGTTTACATATTGGAACAAGGCTTCAACATCTGAACTTATGGATTTTTTTT
HC2A	ACAATATCTGAAGTCTGCCTGCACCAGTTCCAGTACATGGGGAAGCGATACATAGCCAGG
HC2-80	ACAATATCTGAAGTCTGCCTGCACCAGTTCCAGTACATGGGGAAGCGATACATAGCCAGG
HC2B	ACAATATCTGAAGTCTGCCTGCACCAGTTCCAGTACATGGGGAAGCGATACATAGCCAGG
HC2C	-----
HC2D-KIAA1058	ACAATATCTGAAGTCTGCCTGCACCAGTTCCAGTACATGGGGAAGCGATACATAGCCAG-
HC2E	ACAATATCTGAAGTCTGCCTGCACCAGTTCCAGTACATGGGGAAGCGATACATAGCCAGG
HC2F	ACAATATCTGAAGTCTGCCTGCACCAGTTCCAGTACATGGGGAAGCGATACATAGCCAG-

Fig. 3A (cont.)



HC2A	AACCAGGAGGGGTTGGGACCCATAGTTCATGATCGAAAGTCTCAGACATTGCCTGTTTCC
HC2-80	AACCAGGAGGGGTTGGGACCCATAGTTCATGATCGAAAGTCTCAGACATTGCCTGTTTCC
HC2B	AACCAGGAGGGGTTGGGACCCATAGTTCATGATCGAAAGTCTCAGACATTGCCTGTTTCC
HC2C	-----
HC2D-KIAA1058	-----AA-----
HC2E	AACCAGGAGGGGTTGGGACCCATAGTTCATGATCGAAAGTCTCAGACATTGCCTGTTTCC
HC2F	-----TGTGA-----GAAAG-----ATATCAAGTGT-----

HC2A	CGTAACAGAACAGGAATGATGCATGCCAGATTGCAGCAGCTGGGCAGCCTGGATAACTCT
HC2-80	CGTAACAGAACAGGAATGATGCATGCCAGATTGCAGCAGCTGGGCAGCCTGGATAACTCT
HC2B	CGTAACAGAACAGGAATGATGCATGCCAGATTGCAGCAGCTGGGCAGCCTGGATAACTCT
HC2C	-----
HC2D-KIAA1058	-----CAGGAATGATGCATGCCAGATTGCAGCAGCTGGGCAGCCTGGATAACTCT
HC2E	CGTAACAGAACAGGAATGATGCATGCCAGATTGCAGCAGCTGGGCAGCCTGGATAACTCT
HC2F	-----GCTTGGA-----

HC2A	CTCACTTTTAACCACAGCTATGGCCACTCGGACGCAGATGTTCTGCACCAGTCATTACTT
HC2-80	CTCACTTTTAACCACAGCTATGGCCACTCGGACGCAGATGTTCTGCACCAGTCATTACTT
HC2B	CTCACTTTTAACCACAGCTATGGCCACTCGGACGCAGATGTTCTGCACCAGTCATTACTT
HC2C	-----
HC2D-KIAA1058	CTCACTTTTAACCACAGCTATGGCCACTCGGACGCAGATGTTCTGCACCAGTCATTACTT
HC2E	CTCACTTTTAACCACAGCTATGGCCACTCGGACGCAGATGTTCTGCACCAGTCATTACTT
HC2F	-TTTCTGTAGACAATGGCTATGGCCACTCGGACGCAGATGTTCTGCACCAGTCATTACTT

HC2A	GAAGCCAACATTGCTACTGAGGTTTGCCTGACAGCTCTGGACACGCTTTCTCTATTTACA
HC2-80	GAAGCCAACATTGCTACTGAGGTTTGCCTGACAGCTCTGGACACGCTTTCTCTATTTACA
HC2B	GAAGCCAACATTGCTACTGAGGTTTGCCTGACAGCTCTGGACACGCTTTCTCTATTTACA
HC2C	-----
HC2D-KIAA1058	GAAGCCAACATTGCTACTGAGGTTTGCCTGACAGCTCTGGACACGCTTTCTCTATTTACA
HC2E	GAAGCCAACATTGCTACTGAGGTTTGCCTGACAGCTCTGGACACGCTTTCTCTATTTACA
HC2F	GAAGCCAACATTGCTACTGAGGTTTGCCTGACAGCTCTGGACACGCTTTCTCTATTTACA

HC2A	TTGGCGTTTAAGAACCAGCTCCTGGCCGACCATGGACATAATCCTCTCATGAAAAAAGTT
HC2-80	TTGGCGTTTAAGAACCAGCTCCTGGCCGACCATGGACATAATCCTCTCATGAAAAAAGTT
HC2B	TTGGCGTTTAAG-----CTCCTGGCCGACCATGGACATAATCCTCTCATGAAAAAAGTT
HC2C	-----
HC2D-KIAA1058	TTGGCGTTTAAGAACCAGCTCCTGGCCGACCATGGACATAATCCTCTCATGAAAAAAGTT
HC2E	TTGGCGTTTAAGAACCAGCTCCTGGCCGACCATGGACATAATCCTCTCATGAAAAAAGTT
HC2F	TTGGCGTTTAAGAACCAGCTCCTGGCCGACCATGGACATAATCCTCTCATGAAAAAAGTT

HC2A	TTTGATGTCTACCTGTGTTTTCTTCAAAAACATCAGTCTGAAACGGCTTTAAAAAATGTC
HC2-80	TTTGATGTCTACCTGTGTTTTCTTCAAAAACATCAGTCTGAAACGGCTTTAAAAAATGTC
HC2B	TTTGATGTCTACCTGTGTTTTCTTCAAAAACATCAGTCTGAAACGGCTTTAAAAAATGTC
HC2C	-----
HC2D-KIAA1058	TTTGATGTCTACCTGTGTTTTCTTCAAAAACATCAGTCTGAAACGGCTTTAAAAAATGTC
HC2E	TTTGATGTCTACCTGTGTTTTCTTCAAAAACATCAGTCTGAAACGGCTTTAAAAAATGTC
HC2F	A-----

Fig. 3A (cont.)



0568783742828960  
00E101

HC2A	TTCACTGCCTTAAGGTCCTTAATTTATAAGTTTCCCTCAACATTCTATGAAGGGAGAGCG
HC2-80	TTCACTGCCTTAAGGTCCTTAATTTATAAGTTTCCCTCAACATTCTATGAAGGGAGAGCG
HC2B	TTCACTGCCTTAAGGTCCTTAATTTATAAGTTTCCCTCAACATTCTATGAAGGGAGAGCG
HC2C	-----
HC2D-KIAA1058	TTCACTGCCTTAAGGTCCTTAATTTATAAGTTTCCCTCAACATTCTATGAAGGGAGAGCG
HC2E	TTCACTGCCTTAAGGTCCTTAATTTATAAGTTTCCCTCAACATTCTATGAAGGGAGAGCG
HC2F	-----
HC2A	GACATGTGTGCGGCTCTGTGTTACGAGATTCTCAAGTGCTGTAAGTCCAAGCTGAGCTCC
HC2-80	GACATGTGTGCGGCTCTGTGTTACGAGATTCTCAAGTGCTGTAAGTCCAAGCTGAGCTCC
HC2B	GACATGTGTGCGGCTCTGTGTTACGAGATTCTCAAGTGCTGTAAGTCCAAGCTGAGCTCC
HC2C	-----
HC2D-KIAA1058	GACATGTGTGCGGCTCTGTGTTACGAGATTCTCAAGTGCTGTAAGTCCAAGCTGAGCTCC
HC2E	GACATGTGTGCGGCTCTGTGTTACGAGATTCTCAAGTGCTGTAAGTCCAAGCTGAGCTCC
HC2F	-----
HC2A	ATCAGGACGGAGGCCTCCCAGCTGCTCTACTTCCTGATGAGGAACAACCTTTGATTACACT
HC2-80	ATCAGGACGGAGGCCTCCCAGCTGCTCTACTTCCTGATGAGGAACAACCTTTGATTACACT
HC2B	ATCAGGACGGAGGCCTCCCAGCTGCTCTACTTCCTGATGAGGAACAACCTTTGATTACACT
HC2C	-----
HC2D-KIAA1058	ATCAGGACGGAGGCCTCCCAGCTGCTCTACTTCCTGATGAGGAACAACCTTTGATTACACT
HC2E	ATCAGGACGGAGGCCTCCCAGCTGCTCTACTTCCTGATGAGGAACAACCTTTGATTACACT
HC2F	-----
HC2A	GGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATA
HC2-80	GGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATA
HC2B	GGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATA
HC2C	-----
HC2D-KIAA1058	GGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATA
HC2E	GGAAAGAAGTCCTTTGTCCGGACACATTTGCAAGTCATCATATCTGTCAGCCAGCTGATA
HC2F	-----
HC2A	GCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAAC
HC2-80	GCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAAC
HC2B	GCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAAC
HC2C	-----
HC2D-KIAA1058	GCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAAC
HC2E	GCAGACGTTGTTGGCATTGGGGAAACCAGATTCCAGCAGTCCCTGTCCATCATCAACAAC
HC2F	-----
HC2A	TGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTA
HC2-80	TGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTA
HC2B	TGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTA
HC2C	-----
HC2D-KIAA1058	TGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTA
HC2E	TGTGCCAACAGTGACCGGCTTATTAAGCACACCAGCTTCTCCTCTGATGTGAAGGACTTA
HC2F	-----

Fig. 3A (cont.)

09687837  
101300

HC2A	ACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGAC
HC2-80	ACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGAC
HC2B	ACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGAC
HC2C	-----
HC2D-KIAA1058	ACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGAC
HC2E	ACCAAAAGGATACGCACGGTGCTAATGGCCACCGCCCAGATGAAGGAGCATGAGAACGAC
HC2F	-----

HC2A	CCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCCGAG
HC2-80	CCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCCGAG
HC2B	CCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCCGAG
HC2C	-----
HC2D-KIAA1058	CCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCCGAG
HC2E	CCAGAGATGCTGGTGGACCTCCAGTACAGCCTGGCCAAATCCTATGCCAGCACGCCCCGAG
HC2F	-----

HC2A	CTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCA
HC2-80	CTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCA
HC2B	CTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCA
HC2C	-----
HC2D-KIAA1058	CTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCA
HC2E	CTCAGGAAGACGTGGCTCGACAGCATGGCCAGGATCCATGTCAAAAATGGCGATCTCTCA
HC2F	-----

HC2A	GAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAA
HC2-80	GAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAA
HC2B	GAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAA
HC2C	-----
HC2D-KIAA1058	GAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAA
HC2E	GAGGCAGCAATGTGCTATGTCCACGTAACAGCCCTAGTGGCAGAATATCTCACACGGAAA
HC2F	-----

HC2A	G-----
HC2-80	G-----
HC2B	G-----
HC2C	-----
HC2D-KIAA1058	GAAGCAGTCCAGTGGGAGCCGCCCTTCTCCCCACAGCCATAGCGCCTGCCTGAGGAGG
HC2E	G-----
HC2F	-----

HC2A	-----GCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATC
HC2-80	-----GCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATC
HC2B	-----GCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATC
HC2C	-----GTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATC
HC2D-KIAA1058	AGCCGGGGAGGCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATC
HC2E	-----GCGTGTTTAGACAAGGATGCACCGCCTTCAGGGTCATTACCCCAAACATC
HC2F	-----

FIG. 3A (cont.)

05687837 101300

HC2A	GACGAGGAGGCCTCCATGATGGAAGACGTGGGGATGCAGGATGTCCATTTCAACGAGGAT
HC2-80	GACGAGGAGGCCTCCATGATGGAAGACGTGGGGATGCAGGATGTCCATTTCAACGAGGAT
HC2B	GACGAGGAGGCCTCCATGATGGAAGACGTGGGGATGCAGGATGTCCATTTCAACGAGGAT
HC2C	GACGAGGAGGCCTCCATGATGGAAGACGTGGGGATGCAGGATGTCCATTTCAACGAGGAT
HC2D-KIAA1058	GACGAGGAGGCCTCCATGATGGAAGACGTGGGGATGCAGGATGTCCATTTCAACGAGGAT
HC2E	GACGAGGAGGCCTCCATGATGGAAGACGTGGGGA-----
HC2F	-----
HC2A	GTGCTGATGGAGCTCCTTGAGCAGTGCGCAGATGGACTCTGGAAAGCCGAGCGCTACGAG
HC2-80	GTGCTGATGGAGCTCCTTGAGCAGTGCGCAGATGGACTCTGGAAAGCCGAGCGCTACGAG
HC2B	GTGCTGATGGAGCTCCTTGAGCAGTGCGCAGATGGACTCTGGAAAGCCGAGCGCTACGAG
HC2C	GTGCTGATGGAGCTCCTTGAGCAGTGCGCAGATGGACTCTGGAAAGCCGAGCGCTACGAG
HC2D-KIAA1058	GTGCTGATGGAGCTCCTTGAGCAGTGCGCAGATGGACTCTGGAAAGCCGAGCGCTACGAG
HC2E	-----AAGCCGAGCGCTACGAG
HC2F	-----
HC2A	CTCATCGCCGACATCTACAACTTATCATCCCCATTTATGAGAAGCGGAGGGATTT----
HC2-80	CTCATCGCCGACATCTACAACTTATCATCCCCATTTATGAGAAGCGGAGGGATTT----
HC2B	CTCATCGCCGACATCTACAACTTATCATCCCCATTTATGAGAAGCGGAGGGATTTTGAG
HC2C	CTCATCGCCGACATCTACAACTTATCATCCCCATTTATGAGAAGCGGAGGGATTTTGAG
HC2D-KIAA1058	CTCATCGCCGACATCTACAACTTATCATCCCCATTTATGAGAAGCGGAGGGATTTTGAG
HC2E	CTCATCGCCGACATCTACAACTTATCATCCCCATTTATGAGAAGCGGAGGGATTTTGAG
HC2F	-----
HC2A	-----
HC2-80	-----
HC2B	AGGCTGGCCCATCTGTATGACACGCTGCACCGGGCCTACAGCAAAGTGACCGAGGTCATG
HC2C	AGGCTGGCCCATCTGTATGACACGCTGCACCGGGCCTACAGCAAAGTGACCGAGGTCATG
HC2D-KIAA1058	AGGCTGGCCCATCTGTATGACACGCTGCACCGGGCCTACAGCAAAGTGACCGAGGTCATG
HC2E	AGGCTGGCCCATCTGTATGACACGCTGCACCGGGCCTACAGCAAAGTGACCGAGGTCATG
HC2F	-----
HC2A	-----
HC2-80	-----
HC2B	CACTCGGGCCGCAGGCTTCTGGGGACCTACTTCCGGGTAGCCTTCTTCGGGCAGG-----
HC2C	CACTCGGGCCGCAGGCTTCTGGGGACCTACTTCCGGGTAGCCTTCTTCGGGCAGG-----
HC2D-KIAA1058	CACTCGGGCCGCAGGCTTCTGGGGACCTACTTCCGGGTAGCCTTCTTCGGGCAGGCAGCG
HC2E	CACTCGGGCCGCAGGCTTCTGGGGACCTACTTCCGGGTAGCCTTCTTCGGGCAGG-----
HC2F	-----
HC2A	-----CTTTGAAGATGAAGATGGA
HC2-80	-----CTTTGAAGATGAAGATGGA
HC2B	-----GATTCTTTGAAGATGAAGATGGA
HC2C	-----GATTCTTTGAAGATGAAGATGGA
HC2D-KIAA1058	CAATACCAGTTTACAGACAGTGAAACAGATGTGGAGGGATTCTTTGAAGATGAAGATGGA
HC2E	-----GATTCTTTGAAGATGAAGATGGA
HC2F	-----

Fig. 3A (cont.)

09687837 101300

HC2A  
HC2-80  
HC2B  
HC2C  
HC2D-KIAA1058  
HC2E  
HC2F

AAGGAGTATATTTACAAGGAACCCAAACTCACACCGCTGTCGGAAATTTCTCAGAGACTC  
AAGGAGTATATTTACAAGGAACCCAAACTCACACCGCTGTCGGAAATTTCTCAGAGACTC  
AAGGAGTATATTTACAAGGAACCCAAACTCACACCGCTGTCGGAAATTTCTCAGAGACTC  
AAGGAGTATATTTACAAGGAACCCAAACTCACACCGCTGTCGGAAATTTCTCAGAGACTC  
AAGGAGTATATTTACAAGGAACCCAAACTCACACCGCTGTCGGAAATTTCTCAGAGACTC  
AAGGAGTATATTTACAAGGAACCCAAACTCACACCGCTGTCGGAAATTTCTCAGAGACTC  
-----

HC2A  
HC2-80  
HC2B  
HC2C  
HC2D-KIAA1058  
HC2E  
HC2F

CTTAAACTGTACTCGGATAAATTTGGTTCTGAAAATGTCAAAATGATACAGGATTCTGGC  
CTTAAACTGTACTCGGATAAATTTGGTTCTGAAAATGTCAAAATGATACAGGATTCTGGC  
CTTAAACTGTACTCGGATAAATTTGGTTCTGAAAATGTCAAAATGATACAGGATTCTGGC  
CTTAAACTGTACTCGGATAAATTTGGTTCTGAAAATGTCAAAATGATACAGGATTCTGGC  
CTTAAACTGTACTCGGATAAATTTGGTTCTGAAAATGTCAAAATGATACAGGATTCTGGC  
CTTAAACTGTACTCGGATAAATTTGGTTCTGAAAATGTCAAAATGATACAGGATTCTGGC  
-----

HC2A  
HC2-80  
HC2B  
HC2C  
HC2D-KIAA1058  
HC2E  
HC2F

AAGGTCAACCCTAAGGATCTGGATTCTAAGTATGCATACATCCAGGTGACTCACGTCATC  
AAGGTCAACCCTAAGGATCTGGATTCTAAGTATGCATACATCCAGGTGACTCACGTCATC  
AAGGTCAACCCTAAGGATCTGGATTCTAAGTATGCATACATCCAGGTGACTCACGTCATC  
AAGGTCAACCCTAAGGATCTGGATTCTAAGTATGCATACATCCAGGTGACTCACGTCATC  
AAGGTCAACCCTAAGGATCTGGATTCTAAGTATGCCTACATCCAGGTGACTCACGTCATC  
AAGGTCAACCCTAAGGATCTGGATTCTAAGTATGCATACATCCAGGTGACTCACGTCATC  
-----

HC2A  
HC2-80  
HC2B  
HC2C  
HC2D-KIAA1058  
HC2E  
HC2F

CCCTTCTTTGACGAAAAAGAGTTGCAAGAAAGGAAAAACAGAGTTTGAGAGATCCCACAAC  
CCCTTCTTTGACGAAAAAGAGTTGCAAGAAAGGAAAAACAGAGTTTGAGAGATCCCACAAC  
CCCTTCTTTGACGAAAAAGAGTTGCAAGAAAGGAAAAACAGAGTTTGAGAGATCCCACAAC  
CCCTTCTTTGACGAAAAAGAGTTGCAAGAAAGGAAAAACAGAGTTTGAGAGATCCCACAAC  
CCCTTCTTTGACGAAAAAGAGTTGCAAGAAAGGAAAAACAGAGTTTGAGAGATCCCACAAC  
CCCTTCTTTGACGAAAAAGAGTTGCAAGAAAGGAAAAACAGAGTTTGAGAGATCCCACAAC  
-----

HC2A  
HC2-80  
HC2B  
HC2C  
HC2D-KIAA1058  
HC2E  
HC2F

ATCCGCCGCTTCATGTTTGAGATGCCATTTACGCAGACCGGGAAGAGGCAGGGCGGGGTG  
ATCCGCCGCTTCATGTTTGAGATGCCATTTACGCAGACCGGGAAGAGGCAGGGCGGGGTG  
ATCCGCCGCTTCATGTTTGAGATGCCATTTACGCAGACCGGGAAGAGGCAGGGCGGGGTG  
ATCCGCCGCTTCATGTTTGAGATGCCATTTACGCAGACCGGGAAGAGGCAGGGCGGGGTG  
ATCCGCCGCTTCATGTTTGAGATGCCATTTACGCAGACCGGGAAGAGGCAGGGCGGGGTG  
ATCCGCCGCTTCATGTTTGAGATGCCATTTACGCAGACCGGGAAGAGGCAGGGCGGGGTG  
-----

HC2A  
HC2-80  
HC2B  
HC2C  
HC2D-KIAA1058  
HC2E  
HC2F

GAAGAGCAGTGCAAACGGCGCACCATCCTGACAGCCATACTGCTTCCCTTATGTGAAG  
GAAGAGCAGTGCAAACGGCGCACCATCCTGACAGCCATACTGCTTCCCTTATGTGAAG  
GAAGAGCAGTGCAAACGGCGCACCATCCTGACAGCCATACTGCTTCCCTTATGTGAAG  
GAAGAGCAGTGCAAACGGCGCACCATCCTGACAGCCATACTGCTTCCCTTATGTGAAG  
GAAGAGCAGTGCAAACGGCGCACCATCCTGACAGCCATACTGCTTCCCTTATGTGAAG  
GAAGAGCAGTGCAAACGGCGCACCATCCTGACAGCCATACTGCTTCCCTTATGTGAAG  
-----

Fig. 3A (cont.)

HC2A	AAGCGCATCCCTGTCATGTACCAGCACCACACTGACCTGAACCCCATCGAGGTGGCCATT
HC2-80	AAGCGCATCCCTGTCATGTACCAGCACCACACTGACCTGAACCCCATCGAGGTGGCCATT
HC2B	AAGCGCATCCCTGTCATGTACCAGCACCACACTGACCTGAACCCCATCGAGGTGGCCATT
HC2C	AAGCGCATCCCTTTTCATGTACCAGCACCACACTGACCTGAACCCCATCGAGGT--CCATT
HC2D-KIAA1058	AAGCGCATCCCTGTCATGTACCAGCACCACACTGACCTGAACCCCATCGAGGTGGCCATT
HC2E	AAGCGCATCCCTGTCATGTACCAGCACCACACTGACCTGAACCCCATCGAGGTGGCCATT
HC2F	-----

HC2A	GACGAGATGAGTAAGAAGGTGGCGGAGCTCCGGCAGCTGTGCTCCTCGGCCGAGGTGGAC
HC2-80	GACGAGATGAGTAAGAAGGTGGCGGAGCTCCGGCAGCTGTGCTCCTCGGCCGAGGTGGAC
HC2B	GACGAGATGAGTAAGAAGGTGGCGGAGCTCCGGCAGCTGTGCTCCTCGGCCGAGGTGGAC
HC2C	GACGAGATGAGTAAGAAGGTGGCGGAGCTCCGGCAGCTGTGCTCCTCGGCCGAGGTGGAC
HC2D-KIAA1058	GACGAGATGAGTAAGAAGGTGGCGGAGCTCCGGCAGCTGTGCTCCTCGGCCGAGGTGGAC
HC2E	GACGAGATGAGTAAGAAGGTGGCGGAGCTCCGGCAGCTGTGCTCCTCGGCCGAGGTGGAC
HC2F	-----

HC2A	ATGATCAAACCTGCAGCTCAAACCTCCAGGGCAGCGTGAGTGTTTCAGGTCAATGCTGGCCCA
HC2-80	ATGATCAAACCTGCAGCTCAAACCTCCAGGGCAGCGTGAGTGTTTCAGGTCAATGCTGGCCCA
HC2B	ATGATCAAACCTGCAGCTCAAACCTCCAGGGCAGCGTGAGTGTTTCAGGTCAATGCTGGCCCA
HC2C	ATGATCAAACCTGCAGCTCAAACCTCCAGGGCAGCGTGAGTGTTTCAGGTCAATGCTGGCCCA
HC2D-KIAA1058	ATGATCAAACCTGCAGCTCAAACCTCCAGGGCAGCGTGAGTGTTTCAGGTCAATGCTGGCCCA
HC2E	ATGATCAAACCTGCAGCTCAAACCTCCAGGGCAGCGTGAGTGTTTCAGGTCAATGCTGGCCCA
HC2F	-----

HC2A	CTAGCATATGCGCGAGCTTTCTTAGATGATACAAACACAAAGCGATATCCTGACAATAAA
HC2-80	CTAGCATATGCGCGAGCTTTCTTAGATGATACAAACACAAAGCGATATCCTGACAATAAA
HC2B	CTAGCATATGCGCGAGCTTTCTTAGATGATACAAACACAAAGCGATATCCTGACAATAAA
HC2C	CTAGCATATGCGCGAGCTTTCTTAGATGATACAAACACAAAGCGATATCCTGACAATAAA
HC2D-KIAA1058	CTAGCATATGCGCGAGCTTTCTTAGATGATACAAACACAAAGCGATATCCTGACAATAAA
HC2E	CTAGCATATGCGCGAGCTTTCTTAGATGATACAAACACAAAGCGATATCCTGACAATAAA
HC2F	-----

HC2A	GTGAAGCTGCTTAAGGAAGTTTTTCAGGCAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCG
HC2-80	GTGAAGCTGCTTAAGGAAGTTTTTCAGGCAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCG
HC2B	GTGAAGCTGCTTAAGGAAGTTTTTCAGGCAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCG
HC2C	GTGAAGCTGCTTAAGGAAGTTTTTCAGGCAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCG
HC2D-KIAA1058	GTGAAGCTGCTTAAGGAAGTTTTTCAGGCAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCG
HC2E	GTGAAGCTGCTTAAGGAAGTTTTTCAGGCAATTTGTGGAAGCTTGCGGTCAAGCCTTAGCG
HC2F	-----

HC2A	GTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAAC
HC2-80	GTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAAC
HC2B	GTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAAC
HC2C	GTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAAC
HC2D-KIAA1058	GTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAAC
HC2E	GTAAACGAACGTCTGATTAAAGAAGACCAGCTCGAGTATCAGGAAGAAATGAAAGCCAAC
HC2F	-----

FIG. 3A (cont.)



HC2A	TACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAGCAG-----ATCTGCC
HC2-80	TACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAGCAG-----ATCTGCC
HC2B	TACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAGCAG-----ATCTGCC
HC2C	TACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAGCAG-----ATCTGCC
HC2D-KIAA1058	TACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAGCAGCTGGGATGATCTGCC
HC2E	TACAGGGAAATGGCGAAGGAGCTTTCTGAAATCATGCATGAGCAG-----ATCTGCC
HC2F	-----

HC2A	CCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTG
HC2-80	CCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTG
HC2B	CCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTG
HC2C	CCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTG
HC2D-KIAA1058	CCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTG
HC2E	CCCTGGAGGAGAAGACGAGCGTCTTACCGAATTCCCTTCACATCTTCAACGCCATCAGTG
HC2F	-----

HC2A	GGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGATTAC
HC2-80	GGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGATTAC
HC2B	GGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGA----
HC2C	GGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGA----
HC2D-KIAA1058	GGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGATTAC
HC2E	GGACTCCAACAAGCACAATGGTTCACGGGATGACCAGCTCGTCTTCGGTCGTGTGA----
HC2F	-----

HC2A	ATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTGATTTGCAAACCTCAGGATGCTTTCCAA
HC2-80	ATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTGATTTGCAAACCTCAGGATGCTTTCCAA
HC2B	-----
HC2C	-----
HC2D-KIAA1058	ATCTCATGGCCCGTGTGTGGGGACTTGCTTTGTGATTTGCAAACCTCAGGATGCTTTCCAA
HC2E	-----
HC2F	-----

HC2A	AGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAA
HC2-80	AGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCAAGGGGAAGGGGAGAGAAAGGAAA
HC2B	-----
HC2C	-----
HC2D-KIAA1058	AGCCAATCACTGGGGAGACCGAGCACAGGGAGGACCA-GGGGAAGGGGAGAGAAAGGAAA
HC2E	-----
HC2F	-----

HC2A	TAAAGAACAACGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGGTGCACATAT
HC2-80	TAAAGAACAACGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGGTGCACATAT
HC2B	-----
HC2C	-----
HC2D-KIAA1058	TAAAGAACAACGTTATTTCTTAACAGACTTTCTATAGGAGTTGTAAGAAGGTGCACATAT
HC2E	-----
HC2F	-----

Fig. 3A (cont.)



HC2A	TTTTTTAAATCTCACTGGCAATATTCAAAGTTTTTCATTGTGTCTTAACAAAGGTGTGGTA
HC2-80	TTTTTTAAATCTCACTGGCAATATTCAAAGTTTTTCATTGTGTCTTAACAAAGGTGTGGTA
HC2B	-----
HC2C	-----
HC2D-KIAA1058	TTTTTTAAATCTCACTGGCAATATTCAAAGTTTTTCATTGTGTCTTAACAAAGGTGTGGTA
HC2E	-----
HC2F	-----

HC2A	GACACTCTTGAGCTGGACTTAGATTTTATTCTTCCTTGCAGAGTAGTGTTAGAATAGATG
HC2-80	GACACTCTTGAGCTGGACTTAGATTTTATTCTTCCTTGCAGAGTAGTGTTAGAATAGATG
HC2B	-----
HC2C	-----
HC2D-KIAA1058	GACACTCTTGAGCTGGACTTAGATTTTATTCTTCCTTGCAGAGTAGTGTTAGAATAGATG
HC2E	-----
HC2F	-----

HC2A	GCCTACAGAAAAAAAGGTTCTGGGATCTACATGGCAGGGAGGGCTGCACTGACATTGAT
HC2-80	GCCTACAGAAAAAAAGGTTCTGGGATCTACATGGCAGGGAGGGCTGCACTGACATTGAT
HC2B	-----
HC2C	-----
HC2D-KIAA1058	GCCTACAGAAAAAAAGGTTCTGGGATCTACATGGCAGGGAGGGCTGCACTGACATTGAT
HC2E	-----
HC2F	-----

HC2A	GCCTGGGGGACCTTTTGCCTCGACTCGTGCCGGAATCTGATCGTAATCAGGGTACAGAA
HC2-80	GCCTGGGGGACCTTTTGCCTCGACTCGTGCCGGAATCTGATCGTAATCAGGGTACAGAA
HC2B	-----
HC2C	-----
HC2D-KIAA1058	GCCTGGGGGACCTTTTGCCTCGAGGCTGAGCTGGAAAATCTTGAAAATATTTTTT---T
HC2E	-----
HC2F	-----

HC2A	CTTACTAGTTTTGTCTAGGAGTATGTTGTATGACTAGGATTTGTGCTATTATCTCATTCA
HC2-80	CTTACTAGTTTTGTCTAGGAGTATGTTGTATGACTAGGATTTGTGCTATTATCTCATTCA
HC2B	-----
HC2C	-----
HC2D-KIAA1058	TTTCCTGTGGCACATTCAGGTTGAATACAAGAACTATTTTTGTGACTAGTTTTTGATGAC
HC2E	-----
HC2F	-----

HC2A	ACAACATAGAGCAAGAATAGTGAGCTAACTGAGCTAGACACTCAATTAATCCGCTACTGG
HC2-80	ACAACATAGAGCAAGAATAGTGAGCTAACTGAGCTAGACACTCAATTAATCCGCTACTGG
HC2B	-----
HC2C	-----
HC2D-KIAA1058	CTAAGGGAACTGACCATTGTAATTTTTGTACCAGTGAACCAGGAGATTTAGTGCTTTTAT
HC2E	-----
HC2F	-----

FIG. 3A (cont.)

HC2A	CTTCAAGTCAGAACTTTGTCATTAATCATCGACTCCGGGACGGTCATATATGTATTACAT
HC2-80	CTTCAAGTCAGAACTTTGTCATTAATCATCGACTCCGGGACGGTCATATATGTATTACAT
HC2B	-----
HC2C	-----
HC2D-KIAA1058	ATTCATTTTCCTTGCATTTAAGAAAATATGAAAGCTTAAGGAATTATGTGAGCTTAAAACT
HC2E	-----
HC2F	-----

HC2A	TTCTACATTTTAAATACTCACATGGGCTTATGCATTAAGTTTAATTGTGATAAATTTGTG
HC2-80	TTCTACATTTTAAATACTCACATGGGCTTATGCATTAAGTTTAATTGTGATAAATTTGTG
HC2B	-----
HC2C	-----
HC2D-KIAA1058	AGTCAAGCAGTTTAGAACCAAAGGCCTATATTAATAACCGCAACTATGCTGAAAAGTACA
HC2E	-----
HC2F	-----

HC2A	CTGGTCCAGTATATGCAATACACTTTAATGGTTTATTCTTGTCATAAAAATGTGCAATAT
HC2-80	CTGGTCCAGTATATGCAATACACTTTAATGGTTTATTCTTGTCATAAAAATGTGCAATAT
HC2B	-----
HC2C	-----
HC2D-KIAA1058	AAGTAGTACAGTATATTGTTATGTACATATCATTGTTAATACAGTCCTGGCATTCTGTAC
HC2E	-----
HC2F	-----

HC2A	GGAGATGTATACAAGTCTTTACT-----
HC2-80	GGAGATGTATACAAGTCTTTACT-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	ATATATGTATTACATTTCTACATTTTAAATACTCACATGGGCTTATGCATTAAGTTTAAT
HC2E	-----
HC2F	-----

HC2A	-----
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	TGTGATAAATTTGTGCTGTTCCAGTATATGCAATACACTTTAATGTTTTATTCTTGTACA
HC2E	-----
HC2F	-----

HC2A	-----
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	TAAAAATGTGCAATATGGAGATGTATACAGTCTTTACTATATTAGGTTTATAAACAGTTT
HC2E	-----
HC2F	-----

FIG. 3A (cont.)

HC2A	-----
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	TAAGAATTTTCATCCTTTTGCCAAAATGGTGGAGTATGTAATTGGTAAATCATAAATCCTG
HC2E	-----
HC2F	-----

HC2A	-----
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	TGGTGAATGGTGGTGTACTTTAAAGCTGTCACCATGTTATATTTTCTTTTAAGACATTAA
HC2E	-----
HC2F	-----

HC2A	-----
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	TTTAGTAATTTTATATTTGGGAAAATAAAGGTTTTTAATTTTATTTAAGTGAATCACTG
HC2E	-----
HC2F	-----

HC2A	-----
HC2-80	-----
HC2B	-----
HC2C	-----
HC2D-KIAA1058	CCCTGCTGTAATTAAACATTCTGTACCACATCTGTATTAAAAAGACATTGCTGACC
HC2E	-----
HC2F	-----

FIG. 3A (cont.)

HC2A	-----
HC2A-80	-----
HC2B	-----
HC2C	-----
HC2D	ASGNLDKNARFSAIYRQDSNKLSNDDMLKLLADFRKPEKMAKLPVILGNLDITIDNVSSD
HC2E	-----
HC2F	-----

HC2A	-----
HC2A-80	-----
HC2B	-----
HC2C	-----
HC2D	FPNYVNSSYIPTKQFETCSKTPITFEVEEFVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQ
HC2E	-----
HC2F	-----

HC2A	-----VLHHHQNPFEFYDEIK
HC2A-80	-----
HC2B	-----
HC2C	-----
HC2D	KSFAKARNIAICIEFKDSDEEDSQPLKCIYGRPGGPVFTRSAFAAVLHHHQNPFEFYDEIK
HC2E	-----
HC2F	-----

HC2A	IELPTQLHEKHHLTFFHVSCDNSSKGSTKKRDVVETQVGYSWLPLLKDGRVVTSEQHI
HC2A-80	-----
HC2B	-----
HC2C	-----
HC2D	IELPTQLHEKHHLTFFHVSCDNSSKGSTKKRDVVETQVGYSWLPLLKDGRVVTSEQHI
HC2E	-----
HC2F	-----

HC2A	PVSANLPSGYLGYQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYC
HC2A-80	-----
HC2B	-----
HC2C	-----
HC2D	PVSANLPSGYLGYQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYC
HC2E	-----
HC2F	-----

HC2A	QKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLTRATQEEVAVNVTRVI
HC2A-80	-----
HC2B	-----AMEGHVMIAFLPTILNQLFRVLTRATQEEVAVNVTRVI
HC2C	-----
HC2D	QKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLTRATQEEVAVNVTRVI
HC2E	-----AMEGHVMIAFLPTILNQLFRVLTRATQEEVAVNVTRVI
HC2F	-----

FIG. 3B

HC2A	IHVVAQCHEEGLESHLRSYVKYAYKAEPYVASEYKTVHEELTKSMTTILKPSADFLTSNK
HC2A-80	-----
HC2B	IHVVAQCHEEGLESHLRSYVKYAYKAEPYVASEYKTVHEELTKSMTTILKPSADFLTSNK
HC2C	-----
HC2D	IHVVAQCHEEGLESHLRSYVKYAYKAEPYVASEYKTVHEELTKSMTTILKPSADFLTSNK
HC2E	IHVVAQCHEEGLESHLRSYVKYAYKAEPYVASEYKTVHEELTKSMTTILKPSADFLTSNK
HC2F	-----
HC2A	LLRYSWFFFDVLIKSMAQHLENSKVKLLRNQRFPPASYHHAAETVVNMLMPHITQKFGDN
HC2A-80	-----
HC2B	LLRYSWFFFDVLIKSMAQHLENSKVKLLRNQRFPPASYHHAAETVVNMLMPHITQKFGDN
HC2C	-----
HC2D	LLKYSWFFFDVLIKSMAQHLENSKVKLLRNQRFPPASYHHAAETVVNMLMPHITQKFRDN
HC2E	LLRYSWFFFDVLIKSMAQHLENSKVKLLRNQRFPPASYHHAAETVVNMLMPHITQKFGDN
HC2F	-----
HC2A	PEASKNANHSLAVFIKRCFTFMDRGFVFKQINNYISCFAPGDPKTLFEYKFEFLRVVCNH
HC2A-80	-----
HC2B	PEASKNANHSLAVFIKRCFTFMDRGFVFKQINNYISCFAPGDPKTLFEYKFEFLRVVCNH
HC2C	-----
HC2D	PEASKNANHSLAVFIKRCFTFMDRGFVFKQINNYISCFAPGDPKTLFEYKFEFLRVVCNH
HC2E	PEASKNANHSLAVFIKRCFTFMDRGFVFKQINNYISCFAPGDPKTLFEYKFEFLRVVCNH
HC2F	-----
HC2A	EHYIPLNLPMPFGKGRIQRYQDLQDYSLTDEFCRNHFLVGLLLREVGTTALQEFREVRLI
HC2A-80	-----QLDYSLTDEFCRNHFLVGLLLREVGTTALQEFREVRLI
HC2B	EHYIPLNLPMPFGKGRIQRYQDLQDYSLTDEFCRNHFLVGLLLREVGTTALQEFREVRLI
HC2C	-----
HC2D	EHYIPLNLPMPFGKGRIQRYQDLQDYSLTDEFCRNHFLVGLLLREVGTTALQEFREVRLI
HC2E	EHYIPLNLPMPFGKGRIQRYQDLQDYSLTDEFCRNHFLVGLLLREVGTTALQEFREVRLI
HC2F	-----
HC2A	AISVLKNLLIKHSFDDRYASRSHQARIATLYLPLFGLLIENVQRINVRDVSPFPVNAGMT
HC2A-80	AISVLKNLLIKHSFDDRYASRSHQARIATLYLPLFGLLIENVQRINVRDVSPFPVNAGMT
HC2B	AISVLKNLLIKHSFDDRYASRSHQARIATLYLPLFGLLIENVQRINVRDVSPFPVNAGMT
HC2C	-----
HC2D	AISVLKNLLIKHSFDDRYASRSHQARIATLYLPLFGLLIENVQRINVRDVSPFPVNAGMT
HC2E	AISVLKNLLIKHSFDDRYASRSHQARIATLYLPLFGLLIENVQRINVRDVSPFPVNAGMT
HC2F	-----
HC2A	VKDESLALPAVNPLVTPQKGSTLDNSLHKDLLGAISGIASPYTTSTPNINSVRNADSRGS
HC2A-80	VKDESLALPAVNPLVTPQKGSTLDNSLHKDLLGAISGIASPYTTSTPNINSVRNADSRGS
HC2B	VKDESLALPAVNPLVTPQKGSTLDNSLHKDLLGAISGIASPYTTSTPNINSVRNADSRGS
HC2C	-----
HC2D	VKDESLALPAVNPLVTPQKGSTLDNSLHKDLLGAISGIASPYTTSTPNINSVRNADSRGS
HC2E	VKDESLALPAVNPLVTPQKGSTLDNSLHKDLLGAISGIASPYTTSTPNINSVRNADSRGS
HC2F	-----ADSRGS

HC2A	LISTDSGNSLPERNSEKSNSLDKHQSSSTLGNSVVRCDKLDQSEIKSLLMCFLYILKSMS
HC2A-80	LISTDSGNSLPERNSEKSNSLDKHQSSSTLGNSVVRCDKLDQSEIKSLLMCFLYILKSMS
HC2B	LISTDSGNSLPERNSEKSNSLDKHQSSSTLGNSVVRCDKLDQSEIKSLLMCFLYILKSMS
HC2C	-----
HC2D	LISTDSGNSLPERNSEKSNSLDKHQSSSTLGNSVVRCDKLDQSEIKSLLMCFLYILKSMS
HC2E	LISTDSGNSLPERNSEKSNSLDKHQSSSTLGNSVVRCDKLDQSEIKSLLMCFLYILKSMS
HC2F	LISTDSGNSLPERNSEKSNSLDKHQSSSTLGNSVVRCDKLDQSEIKSLLMCFLYILKSMS

HC2A	DDALFTYWNKASTSELMDFFTISEVCLHQFQYMGKRYIARNQEGLGPIVHDRKSQTLPLVS
HC2A-80	DDALFTYWNKASTSELMDFFTISEVCLHQFQYMGKRYIARNQEGLGPIVHDRKSQTLPLVS
HC2B	DDALFTYWNKASTSELMDFFTISEVCLHQFQYMGKRYIARNQEGLGPIVHDRKSQTLPLVS
HC2C	-----
HC2D	DDALFTYWNKASTSELMDFFTISEVCLHQFQYMGKRYIAR-----
HC2E	DDALFTYWNKASTSELMDFFTISEVCLHQFQYMGKRYIARNQEGLGPIVHDRKSQTLPLVS
HC2F	DDALFTYWNKASTSELMDFFTISEVCLHQFQYMGKRYIAS-----VR--KISSVLGIS

HC2A	RNRTGMMHARLQQLGSLDNSLTFNHSYGHSDADVLHQSLLEANIATEVCLTALDTLSLFT
HC2A-80	RNRTGMMHARLQQLGSLDNSLTFNHSYGHSDADVLHQSLLEANIATEVCLTALDTLSLFT
HC2B	RNRTGMMHARLQQLGSLDNSLTFNHSYGHSDADVLHQSLLEANIATEVCLTALDTLSLFT
HC2C	-----
HC2D	---TGMMHARLQQLGSLDNSLTFNHSYGHSDADVLHQSLLEANIATEVCLTALDTLSLFT
HC2E	RNRTGMMHARLQQLGSLDNSLTFNHSYGHSDADVLHQSLLEANIATEVCLTALDTLSLFT
HC2F	V-----D-NG-----YGHSDADVLHQSLLEANIATEVCLTALDTLSLFT

HC2A	LAFKNQLLADHGHNPMLMKKVFVDVYLCFLQKHQSETALKNVFTALRS LIYKFPSTFYEGRA
HC2A-80	LAFKNQLLADHGHNPMLMKKVFVDVYLCFLQKHQSETALKNVFTALRS LIYKFPSTFYEGRA
HC2B	LAFK--LLADHGHNPMLMKKVFVDVYLCFLQKHQSETALKNVFTALRS LIYKFPSTFYEGRA
HC2C	-----
HC2D	LAFKNQLLADHGHNPMLMKKVFVDVYLCFLQKHQSETALKNVFTALRS LIYKFPSTFYEGRA
HC2E	LAFKNQLLADHGHNPMLMKKVFVDVYLCFLQKHQSETALKNVFTALRS LIYKFPSTFYEGRA
HC2F	LAFKNQLLADHGHNPMLMKK-----

HC2A	DMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTHLQVIISVSQLI
HC2A-80	DMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTHLQVIISVSQLI
HC2B	DMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTHLQVIISVSQLI
HC2C	-----
HC2D	DMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTHLQVIISVSQLI
HC2E	DMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTHLQVIISVSQLI
HC2F	-----

HC2A	ADVVGIGETRFOQSLSI INNCANS DRLIKHTSFSSDVKDLTKRIRTVLMATAQMKEHEND
HC2A-80	ADVVGIGETRFOQSLSI INNCANS DRLIKHTSFSSDVKDLTKRIRTVLMATAQMKEHEND
HC2B	ADVVGIGETRFOQSLSI INNCANS DRLIKHTSFSSDVKDLTKRIRTVLMATAQMKEHEND
HC2C	-----
HC2D	ADVVGIGGTRFOQSLSI INNCANS DRLIKHTSFSSDVKDLTKRIRTVLMATAQMKEHEND
HC2E	ADVVGIGETRFOQSLSI INNCANS DRLIKHTSFSSDVKDLTKRIRTVLMATAQMKEHEND
HC2F	-----

HC2A	PEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV TALVAEYLTRK
HC2A-80	PEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV TALVAEYLTRK
HC2B	PEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV TALVAEYLTRK
HC2C	-----
HC2D	PEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV TALVAEYLTRK
HC2E	PEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV TALVAEYLTRK
HC2F	-----

HC2A	-----GVFRQGCTAFRVITPNI DEEASMMEDVGMQDVHFNE
HC2A-80	-----GVFRQGCTAFRVITPNI DEEASMMEDVGMQDVHFNE
HC2B	-----GVFRQGCTAFRVITPNI DEEASMMEDVGMQDVHFNE
HC2C	-----FRQGCTAFRVITPNI DEEASMMEDVGMQDVHFNE
HC2D	EAVQWEPPLLPHSHSACLRRSRGGVFRQGCTAFRVITPNI DEEASMMEDVGMQDVHFNE
HC2E	-----GVFRQGCTAFRVITPNI DEEASMMEDVG-----
HC2F	-----

HC2A	DVLMELLEQCADGLWKAERYELIADIYKLI IPIYEKRR-----
HC2A-80	DVLMELLEQCADGLWKAERYELIADIYKLI IPIYEKRR-----
HC2B	DVLMELLEQCADGLWKAERYELIADIYKLI IPIYEKRRDFERLAHLYDTLHRAYSK
HC2C	DVLMELLEQCADGLWKAERYELIADIYKLI IPIYEKRRDFERLAHLYDTLHRAYSK
HC2D	DVLMELLEQCADGLWKAERYELIADIYKLI IPIYEKRRDFERLAHLYDTLHRAYSK
HC2E	-----KAERYELIADIYKLI IPIYEKRRDFERLAHLYDTLHRAYSK
HC2F	-----

HC2A	-----DFFEDEDGKEYIYKEPKLTPLSE
HC2A-80	-----DFFEDEDGKEYIYKEPKLTPLSE
HC2B	VTEVMHSGRLLGTYFRVAFFGQ-----GFFEDEDGKEYIYKEPKLTPLSE
HC2C	VTEVMHSGRLLGTYFRVAFFGQ-----GFFEDEDGKEYIYKEPKLTPLSE
HC2D	VTEVMHSGRLLGTYFRVAFFGQAQYQFTDSETDVEGFFEDEDGKEYIYKEPKLTPLSE
HC2E	VTEVMHSGRLLGTYFRVAFFGQ-----GFFEDEDGKEYIYKEPKLTPLSE
HC2F	-----

HC2A	ISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPFFDEKELQERKTEF
HC2A-80	ISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPFFDEKELQERKTEF
HC2B	ISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPFFDEKELQERKTEF
HC2C	ISQRLLKLYSDKFGSENVKMTQDSGKVNPKDLDSKYAYIQVTHVIPFFDEKELQERKTEF
HC2D	ISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPFFDEKELQERKTEF
HC2E	ISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYAYIQVTHVIPFFDEKELQERKTEF
HC2F	-----

HC2A	ERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPVMYQHHTDLNP
HC2A-80	ERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPVMYQHHTDLNP
HC2B	ERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPVMYQHHTDLNP
HC2C	ERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPFMYQHHTDLNP
HC2D	ERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPVMYQHHTDLNP
HC2E	ERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTAIHCFPYVKKRIPVMYQHHTDLNP
HC2F	-----





**PBL**  
**lung**  
**placenta**  
**sm intestine**  
**liver**  
**kidney**  
**spleen**  
**thymus**  
**colon**  
**skel muscle**  
**heart**  
**brain**

4

~ 7.5 kb —

FIG. 4B

HC2A	-----
KIAA	ASGNLDKNARFSAIYRQDSNKLSNDDMLKLLADFRKPEKMAKLPVILGNLDITIDNVSSD
rat	-----
HC4	-----
HC1	-----
HC3	-----
HC5	-----

HC2A	-----
KIAA	FPNYVNSSYIPTKQFETCSKTPITFEVEEFVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQ
rat	-----
HC4	-----
HC1	-----
HC3	-----
HC5	-----

HC2A	-----VLHHHQNPFEFYDEIK
KIAA	KSFAKARNIAICIEFKDSDEEDSQPLKCIYGRPGGPVFTRSAFAAVLHHHQNPFEFYDEIK
rat	-----
HC4	-----
HC1	-----
HC3	-----
HC5	-----

HC2A	IELPTQLHEKHHLTFFHVSCDNSSKGSTKKRDVVETQVGYSWLPLLKDGRVVTSEQHI
KIAA	IELPTQLHEKHHLTFFHVSCDNSSKGSTKKRDVVETQVGYSWLPLLKDGRVVTSEQHI
rat	-----
HC4	-----
HC1	-----
HC3	-----
HC5	-----

HC2A	PVSANLPSGYLGYQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYC
KIAA	PVSANLPSGYLGYQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYC
rat	-----
HC4	-----
HC1	-----
HC3	-----GPGPARSTVSIISLISNSARV
HC5	-----

HC2A	OKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLT-RATQEEVAVNVTRV
KIAA	OKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLT-RATQEEVAVNVTRV
rat	-----
HC4	-----MEIQVLIRFLSVILMQLFWVLPNMIHEDDVPISCPMV
HC1	-----MSFLPIILNQLFKVLV-QNEEDEITTTVTRV
HC3	NRSRSLSNSNPDISGTPTSPDDEVRSIIGSKGLDRSNSWVNTGGPKAAPWGSPSPSAES
HC5	-----

FIG. 5A

HC2A	I IHVVAQCHEE	SHLRSYVKYAYKAEPYVASEYKTVHEEL	TTILKPSADFLTSN
KIAA	I IHVVAQCHEE	SHLRSYVKYAYKAEPYVASEYKTVHEEL	TTILKPSADFLTSN
rat	-----		
HC4	LFHIVSKCHEEGLDSYLSSFIKYSFRPGKPSAPQAPLIHETLATMMIALLKQSADFLAIN		
HC1	LPDIVAKCHEEQLDHSVQSYIKFVFKTR---	ACKERPVEDHLAKNVTGLLK-SNDSPTVK	
HC3	TQAMDRSCNRMSSHTETSSFLQTLTGRLP----	TKKLFHEELALQWVCSG--SVR---	E
HC5	-----		

Cadherin  
Cleavage

HC2A	KLLRYSWFFFDVLIKSMAQH LIENSKVKLLRNQRF	PASYHHAETVVMMLMPHITQKFGD
KIAA	KLLKYSWFFFDVLIKSMAQH LIENSKVKLLRNQRF	PASYHHAETVVMMLMPHITQKFRD
rat	-----	
HC4	KLLKYSWFFFEIIAKSMATYLLEENKIKLTHGQRF	PKAYHHALHSLFLAIT-IVESQYAE
HC1	HVLKHSWFFFAIILKSMAQH LIDTNKIQLPRPQRF	PESYQNELDNLMVLSDHVIWKYKD
HC3	SALQQAWFFFEIMVKSMVHHLYFNDKLEAPRKS	RFPERFMD DIAALVSTIASDIVSRFQK
HC5	-----	

HC2A	NPEASKNANHSLAVFIKRCFTFMDRGFVFKQIN---	NYIS--CFAPGDPKTLFEYKFEFL
KIAA	NPEASKNANHSLAVFIKRCFTFMDRGFVFKQIN---	NYIS--CFAPGDPKTLFEYKFEFL
rat	-----	
HC4	IPKESRNVNYSLASFLKCCLTLMDRGFVFNLIN---	DYIS--GFSPKDPKVLAEYKFEFL
HC1	ALEETRRATHSVARFLKRCFTFMDRGCVFKMVN---	NYIS--MFSSGDLKTL CQYKFDL
HC3	DTEMVERLNTSLAFLNDLLSVMDRGFVFSLIKSCYQVSSKLYSLPNPSVLVSLRLDFL	
HC5	-----	

HC2A	RVVCNHEHYIPLNLPM-----	PFGKGRIQR-----	YQDLQL----	DYSLTDEF
KIAA	RVVCNHEHYIPLNLPM-----	PFGKGRIQR-----	YQDLQL----	DYSLTDEF
rat	-----			
HC4	QTICNHEHYIPLNLPM-----	AFAPKPKLQR-----	VQDSNL----	EYSLSDEY
HC1	QEVCOHEHFIPCLPIRSANIPDPLTPSES-----		TQELHASDMPEYSVTNEF	
HC3	RIICSHEHYVTNLNLPCLLTPASPSPSVSSATSQSSGFSTNVQDQKIANMFELS--	VPF		
HC5	-----MNADTAPTSPCPSIS--SONSSSCSSFQDQKIASMFDRTSRVPA			

HC2A	CRNHFLVGLLLREVG TALQEFRE----	VRLIAISVLKNLLIKHSFDDRYASRSHQARIAT
KIAA	CRNHFLVGLLLREVG TALQEFRE----	VRLIAISVLKNLLIKHSFDDRYASRSHQARIAT
rat	-----	
HC4	CKHHFLVGLLLRETSIALQDNYE----	IRYTAISVIKNLLIKHAFDTRYQHKNQQAKIAQ
HC1	CRKHFLIGILLREVG FALQEDQD----	VRHLALAVLKNLMAKHSFDDRYREPRKQAQIAS
HC3	RQOHYLAGLVLT ELAVILDPDAEGLFGLHKKVINMVHNLSSHDSDP RYSDPQIKARVAM	
HC5	SSTS-SPGLLFT ELAAALDAEGEGISEVQRKAVSAIHSLLSSHDLDP RCVKPEVKVKIAA	

HC2A	LYLPLFGLLLIENVQRINVRDVSPFPVNAG-MTVKDESALPAVNPLVTPQKGSTLDNSLH	
KIAA	LYLPLFGLLLIENVQRINVRDVSPFPVNAG-MTVKDESALPAVNPLVTPQKGSTLDNSLH	
rat	-----	
HC4	LYLPLFVGLLLIENIQRLAGRDTLYSCAMPNSASRDEFPCG-----	FTSP--AN--RGSLS
HC1	LYMPLYGMLLDNMPRIYLDLYPFTVNTSNQGSRDDLTNGGFQSQTAIKHANSVDTSFS	
HC3	LYLPLIGIIMETVPQLYDFTETHNQGRPIC IATDDYESE-----	SG---SMIS
HC5	LYLPLVGIILDALPQLCDFTVADTRRYR---	TSGSDEEQE-----GA---GAIT

HC2A	KDLLGAISGIASPYTTSTPNINSVRNADSRGSLISTDSGNSLPERNSEKSNSLDKHQQSS	
KIAA	KDLLGAISGIASPYTTSTPNINSVRNADSRGSLISTDSGNSLPERNSEKSNSLDKHQQSS	
rat	-----	
HC4	TDKDTAYGSFQNG-----	HGIKREDSRGSLIP-EGATGFDPDQGN TGEN-----TRQS
HC1	KDVLNSIAAFSS-----	IAISTVNHADSRASLASLDSNPSTNEKSSEKTDNCEKIPRPL
HC3	QTVAMAIAGTSVPQ-----	LTRPGSFLLTSTSGRQHT-----
HC5	QNVALAIAGNNFN-----	LKTSG-IVLSSLPYKQYN-----



HC2A	TLGNSVVRCDKLDQSEIKSLLMCFLYILKSMSDDALFTYWN-KASTSELMDFFTISEVCL
KIAA	TLGNSVVRCDKLDQSEIKSLLMCFLYILKSMSDDALFTYWN-KASTSELMDFFTISEVCL
rat	-----
HC4	STRSSVSQYNRLDQYEIRSLLMCYLYIVKMISEDTLITYWN-KVSPQELINILILLEVCL
HC1	ALIGSTLRFDRLDQAETRSLLMCFLHIMKTISYETLIAYWQ-RAPSPEVSDFFSILDVCL
HC3	-----TFSAESSRSLICLLWVLKN-ADETVLQKWFTDLSVLQLNRLLDLLYLCV
HC5	-----MLNADTTRNLMICFLWIMKN-ADQSLIRKWIADLPSTQLNRILDLLFICV

HC2A	HQFQYMGKRYIARNQEGLG--PIVHDRKS-----QTLPVSRNRTGMM
KIAA	HQFQYMGKRYIAR-----TGMM
rat	-----
HC4	FHFRYMGKRNIARVHDAWLSKHFGIDRKS-----QTMPALRNRSGVM
HC1	QNFYRLGKRNIIRKIAAAF--KFVQSTQNNGTLKGSNPSCQTSGLLAQWMHSTSRHEGHK
HC3	SCFEYKGGKVFERMNSLTFK--KSKDMRAK-----LEEAILGSIGARQEMV
HC5	LCFEYKGGKQSSDKVSTQVLQ--KSRDVKAR-----LEEALLRGEGERGEMM

HC2A	HARLQQL-----GSLDNS-----LTFNHSYGHSDADVLHQSLLEANIATEVC
KIAA	HARLQQL-----GSLDNS-----LTFNHSYGHSDADVLHQSLLEANIATEVC
rat	-----
HC4	QARLQHL-----SSLESS-----FTLNHSSTTTEADIFHQALLEGNTATEVS
HC1	QHRSQTLPIIRGK--NALSNNPKL---LQMLDNTMTSNSNEIDIVHHVDTEANIATEGC
HC3	RRSRGQLERSPSGSAFGSQENLRWRKDMTHWRQNTTEKLDKSRAEIEHEALIDGNLATEAN
HC5	RRRAPGNDRFP-----GLNENLRWKKEQTHWRQANEKLDKTKAELDQEALISGNLATEAH

HC2A	LTALDTLSLFTLAFKNQLLADHGHNPMLKKVFDVYLCFLQKHQSETALKNVFTALRS LIY
KIAA	LTALDTLSLFTLAFKNQLLADHGHNPMLKKVFDVYLCFLQKHQSETALKNVFTALRS LIY
rat	-----KLSRGHSPLMKKVFVDVYLCFLQKHQSEMAKKNVFTALRS LIY
HC4	LTVLDTISFFTQCFKTHFLNNDGHNPMLKKVFDIHLAFLKNGQSEVSLKHVFASLRAFIS
HC1	LTILDVLSLFTQTHQRQLQQCDQNSLMKRGFDTYMLFFQVNOQSATALKHVFASLRLEVC
HC3	LIILDTLEIVVQTVS--VTES--KESILGGVLKVLLHSMACNQSAVYLQHC FATQRALVS
HC5	LIILDMQENIIQASS--ALDC--KDSLLGGVLRVLVNSLNCQSTTYLTHCFATLRALIA

HC2A	KFPSTFYEGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTH
KIAA	KFPSTFYEGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTH
rat	KFPSTFYEGRADMCASLCYEVLKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTH
HC4	KFPSAFFKGRVNMCAAFCEVLKCCCTSKISSSTRNEASALLYLLMRNNFEYTKRKTFLRTH
HC1	KFPSAFFQGPADLCGSFCYEVLKCCNHRSRSTQTEASALLYLFMRKNFEFNKQKSIVRSH
HC3	KFPELLFEEETEQCADLCLRLLRHCSSSIGTIRSHPSASLYLLMRQNF EIGN--NFARVK
HC5	KFGDLLFEEEEVEQC FDLCHQVLHHCSSSMDVTRSQACATLYLLMRFSFGATS--NFARVK

HC2A	LQV IISVSQLIADVVGIGETR FQQSLSI INNCANS DR LIKHTSFSSDVKDLTKRI RTVLM
KIAA	LQV IISVSQLIADVVGIGETR FQQSLSI INNCANS DR LIKHTSFSSDVKDLTKRI RTVLM
rat	LQV IISLSQLIADVVGIGETR FQQSLSI INNCANS DR LIKHTSFSSDVKDLTKRI RTVLM
HC4	LQ I I IAVSQLIADVALSGGSRFQESLFI INNFANS DRPMLARAFPAEVKDLTKRI RTVLM
HC1	LQ LIKAVSQLIAD-AGIGGSRFQHS LAITNNFANGDKQMKNSNFP AEVKDLTKRI RTVLM
HC3	MQV PMSLSSLVGT SQNFNEEFLRRSLKTI LTYAEEDLELRETTFPDQVQDLVFNLMILS
HC5	MQVTMSLASLVGRAPDFNEEHLRRSLRTI LAYSEEDTAMQMT PFPTQVEELLCNLNS ILY

HC2A	ATAQMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV
KIAA	ATAQMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV
rat	ATAQMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV
HC4	ATAQMKEHEKDP EMLIDLQYSLAKSYASTPELRKTWLD SMAIHVKN GDFSEAAMCYVHV
HC1	ATAQMKEHEKDP EMLVDLQYSLANSYASTPELRRTWLESMAKIHARNGDLSEAAMCYIHI
HC3	DTVKMKEHQEDPEMLIDL MYRIAKGYQTSPDLRLTWLQNMAGKHSERSNHAEEAAQCLVHS
HC5	DTVKMREFQEDPEMLMDL MYRIAKSYQASPDRLRLTWLQNMAEKHTKKKCYTEAMCLVHA

## SH3

HC2A	TALVAEYI	TRKGV-----	-----	FRQGCTAFRVITPN
KIAA	TALVAEYI	TRKEA-----	VQWEPPLLPHSHSACLRRSRGGVFRQGCTAFRVITPN	
rat	TALVAEYI	TRKEAD-----	LALQREPPVFPYSHTSCQRKSRGGMFRQGCTAFRVITPN	
HC4	AALVAEFI	HRKKL-----	-----	FPNGCSAFKKITPN
HC1	AALIAEYI	KRKG YWKVEKI	CTASLLSEDT HPCDSNSLLTTPSG	GSMFSMGWPAFLSITPN
HC3	AALVAEYI	SMLD-----	-----	RKYL PVGCVTFQNISSN
HC5	AALVAEYI	SMLD-----	-----	HSYL PVGSVSFQNISSN

HC2A	IDEEASMMEDVGMQD-----	VHFNE DVLME LLEQCADGLWKAERYELIADIYKLI IPI
KIAA	IDEEASMMEDVGMQD-----	VHFNE DVLME LLEQCADGLWKAERYELIADIYKLI IPI
rat	IDEEASMMEDVGMQD-----	VHFNE DVLME LLEQCADGLWKAERLRAGLLTSINSSSP
HC4	IDEEGAMKEDAGMMD-----	VHYSEEVLLELLEQCVNGLWKAERYEI ISEISKLIGPI
HC1	IKEEGA AKEDSGMHD-----	TPYNE NILVEQLYMCGEFLWKSE RYELIADV NKP IIAV
HC3	VLEESAVSDDV VSPDEEGICSGKYFTESGLVGLLEQAAASF SMAGMYEAVNEVYKVL IPI	
HC5	VLEESV VSED TLSPDE DGV CAGQYFTESGLVGLLEQAAELFSTGGLYETVNEVYKVL IPI	

		ITAM	ITAM		ITAM	ITAM
HC2A	YEKRRD-----					
KIAA	YEKRRD FERLAHL YDTLH RAYSKVTEVMHSGRLLGTYFRVAFFGQAAQYQFTDSETDVE					
rat	SMKSGGTLETTHLYDTLH RPYSKVTEVITR-----			A-----	AGSWDLLPGGLFGQ	
HC4	YENRREFENLTQVYRTIHGAYTKILEVMHTKKRLLG-----				TFFRVAFYGO	
HC1	FEKQ RDFKKLS DLYYDIHRSYLKVAE VVNSEKRLFG-----				RYFRVAFYGO	
HC3	HEANRDAKKLSTIHGKLQEA FSKIVHQSTGWERMFG-----				TYFRVGFG-	
HC5	LEAHREFRKLTLTHSKLQRA FDSIVNKDH--KRMFG-----				TYFRVGFG-	

		ITAM	ITAM
HC2A	-FFEDEDGKEYIYKEPKLTPLSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYA		
KIAA	GFFEDEDGKEYIYKEPKLTPLSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYA		
rat	GFFEDEDGKEYIYKEPKLTPLSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKFA		
HC4	SFFEEDGKEYIYKEPKLTGLSEISLRLVKLYGEKFGTENVKIIQSDKVNAKELDPKYA		
HC1	GFFEEEGKEYIYKEPKLTGLSEISQRLLKLYADKFGADNVKIIQDSNKVNPKDLDPKYA		
HC3	TKFGDLDEQEFVYKEPAITKLAEISHRLEGEFYGERFGEDVVEVIKDSNPVDKCKLDPNKA		
HC5	SKFGDLDEQEFVYKEPAITKLPEISHRLEAFYGOFGAEFVEVIKDSTPVDKTKLDPNKA		

## ITAM

HC2A	YIQVTHVIPFFDEKELQERKTEFERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTA
KIAA	YIQVTHVIPFFDEKELQERKTEFERSHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTA
rat	YIQVTHVTPFFDEKELQERKTEFERCHNIRRFMFEMPFTQTGKRQGGVEEQCKRRTILTA
HC4	HIQVTYVKPYFDDKELTERKTEFERNHNISR FVFEAPYTL SGKKQGCIEEQCKRRTILT T
HC1	YIQVTYVTPFFEEKEIEDRKTD FEMHNNIR FVFETPFTLSGKKHGGVAEQCKRRTILT T
HC3	YIQITYVEPYFD TYEMKDRITYFDKNYNLRRFMYCTPFTLDGRAHGE LHEQFKRKTILT T
HC5	YIQITFVEPYFDEYEMKDRVTYFEKNFNLRRFMYTTPFTLEGRPRGELHEQYRRNTVLT T

FIG. 5A (cont.)



Coiled-Coil

HC2A	IHC	F	P	Y	V	K	K	R	I	P	V	M	Y	Q	H	H	T	D	I	N	P	I	E	V	A	I	D	E	M	S	K	K	V	A	E	L	R	Q	L	C	S	S	A	E	V	D	M	I	K	L	Q	L	K	L	Q	G	S	V		
KIAA	IHC	F	P	Y	V	K	K	R	I	P	V	M	Y	Q	H	H	T	D	I	N	P	I	E	V	A	I	D	E	M	S	K	K	V	A	E	L	R	Q	L	C	S	S	A	E	V	D	M	I	K	L	Q	L	K	L	Q	G	S	V		
rat	IHC	F	P	Y	V	K	K	R	I	P	V	M	Y	Q	H	H	T	D	I	N	P	I	E	V	A	I	D	E	M	S	K	K	V	A	E	L	H	Q	L	C	S	S	A	E	V	D	M	I	K	L	Q	L	K	L	Q	G	S	V		
HC4	S	N	S	F	P	Y	V	K	K	R	I	P	I	N	C	E	Q	Q	I	N	I	K	P	I	D	G	A	T	D	E	I	K	D	K	T	A	E	L	Q	K	L	C	S	S	T	D	V	D	M	I	Q	L	Q	L	K	L	Q	G	W	V
HC1	S	H	L	F	P	Y	V	K	K	R	I	Q	V	I	S	S	T	E	I	N	P	I	E	V	A	I	D	E	M	S	R	K	V	S	E	L	N	Q	L	C	T	M	E	E	V	D	M	I	S	L	Q	L	K	L	Q	G	S	V		
HC3	S	H	A	F	P	Y	I	K	T	R	V	N	V	T	H	K	E	E	I	I	L	T	P	I	E	V	A	I	E	D	M	Q	K	T	Q	E	L	A	F	A	T	H	Q	D	P	A	D	P	K	M	L	Q	M	V	L	Q	G	S	V	
HC5	M	H	A	F	P	Y	I	K	T	R	I	S	V	I	Q	K	E	E	F	V	L	T	P	I	E	V	A	I	E	D	M	K	K	T	L	Q	L	A	V	A	I	N	Q	E	P	P	D	A	K	M	L	Q	M	V	L	Q	G	S	V	

Coiled-Coil 2

HC2A	S	V	Q	V	N	A	G	P	L	A	Y	A	R	A	F	L	D	D	T	N	T	K	R	Y	P	D	N	K	V	K	L	L	K	E	V	F	R	Q	F	V	E	A	C	G	Q	A	I	A	V	N	E	R	L	I	K	E	D	Q	L	E
KIAA	S	V	Q	V	N	A	G	P	L	A	Y	A	R	A	F	L	D	D	T	N	T	K	R	Y	P	D	N	K	V	K	L	L	K	E	V	F	R	Q	F	V	E	A	C	G	Q	A	I	A	V	N	E	R	L	I	K	E	D	Q	L	E
rat	S	V	Q	V	N	A	G	P	L	A	Y	A	R	A	F	L	D	D	T	N	T	K	R	Y	P	D	N	K	V	K	L	L	K	E	V	F	R	Q	F	V	E	A	C	G	Q	A	I	A	V	N	E	R	L	I	K	E	D	Q	L	E
HC4	S	V	Q	V	N	A	G	P	L	A	Y	A	R	A	F	L	N	D	S	Q	A	S	K	Y	P	P	K	V	S	E	L	K	D	M	F	R	K	F	I	Q	A	C	S	I	A	I	E	L	N	E	R	L	I	K	E	D	Q	V		
HC1	S	V	K	V	N	A	G	P	M	A	Y	A	R	A	F	L	E	E	T	N	A	K	K	Y	P	D	N	Q	V	K	L	L	K	E	I	F	R	Q	F	A	D	A	C	G	Q	A	I	D	V	N	E	R	L	I	K	E	D	Q		
HC3	G	T	T	V	N	Q	G	P	L	E	V	A	Q	V	F	L	S	E	I	P	S	D	P	K	L	F	R	H	H	N	K	L	R	L	C	F	K	D	F	T	K	R	C	E	D	A	I	R	K	N	K	S	L	I	G	P	V	Q		
HC5	G	A	T	V	N	Q	G	P	L	E	V	A	Q	V	F	L	A	E	I	P	A	D	P	K	L	Y	R	H	H	N	K	L	R	L	C	F	K	E	F	I	M	R	C	G	E	A	V	E	K	N	K	R	L	I	T	A				

Coiled-Coil 2

HC2A	Y	Q	E	E	M	K	A	N	Y	R	E	M	A	K	E	L	S	E	I	M	H	E	Q	I	C	P	L	E	E	K	T	S	-	V	L	P	N	S	L	H	I	F	N	A	I	S	G	T	P	T	S	T	M	V	H	G	M	T	S
KIAA	Y	Q	E	E	M	K	A	N	Y	R	E	M	A	K	E	L	S	E	I	M	H	E	Q	L	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-							
rat	Y	Q	E	E	M	K	A	N	Y	R	E	I	R	K	E	L	S	D	I	I	V	E	R	I	C	P	G	E	D	K	R	A	T	K	F	P	A	H	L	Q	R	H	Q	R	D	T	N	K	H	S	G	S	R	V	D	Q	F	I	L
HC4	Y	H	E	G	L	K	S	N	F	R	D	M	V	K	E	L	S	D	I	I	H	E	Q	I	L	Q	E	D	T	M	H	S	P	W	M	S	N	T	L	H	V	F	C	A	I	S	G	T	S	S	D	R	G	Y	G	S	P	R	
HC1	Y	Q	E	E	L	R	S	H	Y	K	D	M	L	S	E	L	S	T	V	M	N	E	Q	I	T	G	R	D	D	L	S	K	-	-	-	R	G	V	D	Q	T	C	T	R	V	I	S	K	A	T	P	A	L	P	T	V			
HC3	Y	Q	R	E	L	G	-	-	-	-	K	L	S	S	-	-	-	-	-	-	-	-	P	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-									
HC5	Y	Q	E	L	K	K	N	Y	N	K	L	K	E	N	L	R	P	M	I	E	R	K	I	P	E	L	Y	K	P	I	F	R	V	E	S	Q	K	R	D	S	F	H	R	S	S	F	R	K	C	E	T	Q	L	S	Q	G			

PBM

HC2A	S	S	V	V	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
KIAA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
rat	C	V	T	L	P	H	E	P	H	V	G	T	C	F	V	M	C	K	L	R	T	T	F	R	A	N	H	W	F	C	Q	A	Q	E	E	A	M	G	N	G	R	E	K	E	P	W	T	V	I	F	N	S	R	F	Y
HC4	E	V	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
HC1	S	A	E	V	Z	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
HC3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
HC5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					

HC2A	-	-	-	-	-
KIAA	-	-	-	-	-
rat	V	H	I	F	F
HC4	-	-	-	-	-
HC1	-	-	-	-	-
HC3	-	-	-	-	-
HC5	-	-	-	-	-

096383 10100

	A	B	
CLASP-1	YRVAFYGQ	EGKEYIYKEP	
KIAA1058	FRVAFFGQ	AAQYQFTDSETDVEGFFEDGKEYIYKEP	
CLASP-2		FEDEDGKEYIYKEP	
CLASP-6	FRVAFFGQ	EGKEYIYKEP	
CLASP-4	FRVAFYGQ	SFFEEEDGKEYIYKEP	
DOCK180	FAVGYYGQ	GFPTFLRGKVFIYRGKEYERRED	
DOCK2	FAVGYYGQ	GFPSFLRNKVFIYRGKEYERRED	
DOCK3	FRVGFYGR	KFPFFLRNKEYVCRGH	
KIAA0716	FRVGFYGR	KFPFFLRNKEFVCRGH	
CLASP-3	FRVGFYGT	KFGDLDEQEFVYKEP	
CONSENSUS	F V FYG YF	KEY K Q F R	
			C
TRG	PKLTPLSEISQRL	LLKLYSDKFGSENVKMIODSGKVNPKDLDSKFAY	IOVTHVTPFFDEKE
CLASP-1	PKLTGLSEISQRL	LLKLYADKFGADNVKIIQDSNKNVNPKDLDPKYAY	IQVTYVTPFFEEKE
CLASP-2	PKLTPLSEISQRL	LLKLYSDKFGSENVKMTQDSGKVNPKDLDSKYAY	IOVTHVIPFFDEKE
CLASP-4	PKLTGLSEISLRL	VKLYGEKFGTENVKIIQSDKVNAKELDPKYAH	IQVTYVKPYFDDKE
CLASP-3	PAITKLAEISHR	LEGFYGERFGEDVVEVIKDSNPVDKCKLDPNKAY	IQITYVEPYFDTYE
KIAA0716	HDYERLEAFQQR	MLNEFPFAIA-----MQHANQPDETI	FQAEAYLQIYAVTPIPESQE
DOCK3	HDYERLEAFQQR	MLSEFPQAVA-----MQHPNHPDDAILQ	CDAYLQIYAVTPIPDPYVD
DOCK2		FQMQLMTQFPNAEK-----MNTTSAPGDDVKN	APGQYIQCFTVQPVLDDEHP
DOCK180	EYERREDFQMQL	MTQFPNAEK-----MNTTSAPGDDVKN	APGQYIQCFTVQPVLDDEHP
CONSENSUS	L L Y M F		YIQ+ V P D L E
	D	E	
CLASP-1	RTILTTSHLFPYV	KKRIQVISQSSTELNPIEVAIDEMSRKVSELN	
TRG	RTILTAIHCFPYV	KKRIPVMYQHHTDLNPIEVAIDEMSKKVAELH	
KIAA1058	RTILTAIHCFPYV	KKRIPVMYQHHTDLNPIEVAIDEMSKKVAELR	
CLASP-2	RTILTAIHCFPYV	KKRIPVMYQHHTDLNPIEVAIDEMSKKVAELR	
CLASP-6	RTILTAIHCFPYV	KKRIPFMYQHHTDLNPIEV:HDEM	SKKVAELR
CLASP-4	RTILTTNSNFPYV	KKRIPINCEQQINLKPIDVATDEIKDKTAELQ	
CLASP-3	KTILTTSHAFPYI	KTRVNVTHKEEIIILTPIEVAIEDMQKKTQELA	
CLASP-5	NTVLTTMHAFPYI	KTRISVIQKEEFVLTPIEVAIEDMKKKTQLA	
KIAA0716	RTSLYLVSQSLPG	ISRWFVEVEKREVVEMSPLENAIENVLENKNQQLK	
DOCK2	RTSFVTAYKLP	GILRWFEVVHMSQTTISPLENAIETMSTANEKIL	
DOCK3	RTTLTLTHSLPG	ISRWFEEVERRELVEVSPLENAIQVVENKNQELR	
DOCK180	RTSFVTAYKLP	GILRWFEVVHMSQTTISPLENAIETMSTANEKIL	
CONSENSUS	RT L F	FP V + V L L	+ P+E AI+ M +L + I

FIG. 5B

[illegible]

CLASP2variant=KIAA1058

FIG. 5B (cont.)

2 32  
GTT TTA CAC CAT CAC CAA AAC CCA GAA TTT TAT GAT GAG ATT AAA ATA GAG TTG CCC ACT  
val leu his his his gln asn pro glu phe tyr asp glu ile lys ile glu leu pro thr

62 92  
CAG CTG CAT GAA AAG CAC CAC CTG TTG CTC ACA TTC TTC CAT GTC AGC TGT GAC AAC TCA  
gln leu his glu lys his his leu leu leu thr phe phe his val ser cys asp asn ser

122 152  
AGT AAA GGA AGC ACG AAG AAG AGG GAT GTC GTT GAA ACC CAA GTT GGC TAC TCC TGG CTT  
ser lys gly ser thr lys lys arg asp val val glu thr gln val gly tyr ser trp leu

182 212  
CCC CTC CTG AAA GAC GGA AGG GTG GTG ACA AGC GAG CAG CAC ATC CCG GTC TCG GCG AAC  
pro leu leu lys asp gly arg val val thr ser glu gln his ile pro val ser ala asn

242 272  
CTT CCT TCG GGC TAT CTT GGC TAC CAA GAG CTT GGG ATG GGC AGG CAT TAT GGT CCG GAA  
leu pro ser gly tyr leu gly tyr gln glu leu gly met gly arg his tyr gly pro glu

302 332  
ATT AAA TGG GTA GAT GGA GGC AAG CCA CTG CTG AAA ATT TCC ACT CAT CTG GTT TCT ACA  
ile lys trp val asp gly gly lys pro leu leu lys ile ser thr his leu val ser thr  
ref 1.1, 1.2 and 1.3

362 392  
GTG TAT ACT CAG GAT CAG CAT TTA CAT AAT TTT TTC CAG TAC TGT CAG AAA ACC GAA TCT  
val tyr thr gln asp gln his leu his asn phe phe gln tyr cys gln lys thr glu ser

422 452  
GGA GCC CAA GCC TTA GGA AAC GAA CTT GTA AAG TAC CTT AAG AGT CTG CAT GCG ATG GAA  
gly ala gln ala leu gly asn glu leu val lys tyr leu lys ser leu his ala met glu

482 512  
GGC CAC GTG ATG ATC GCC TTC TTG CCC ACT ATC CTA AAC CAG CTG TTC CGA GTC CTC ACC  
gly his val met ile ala phe leu pro thr ile leu asn gln leu phe arg val leu thr

542 572  
AGA GCC ACA CAG GAA GAA GTC GCG GTT AAC GTG ACT CGG GTC ATT ATT CAT GTG GTT GCC  
arg ala thr gln glu glu val ala val asn val thr arg val ile ile his val val ala

602 632  
CAG TGC CAT GAG GAA GGA TTG GAG AGC CAC TTG AGG TCA TAT GTT AAG TAC GCG TAT AAG  
gln cys his glu glu gly leu glu ser his leu arg ser tyr val lys tyr ala tyr lys

662 692  
GCT GAG CCA TAT GTT GCC TCT GAA TAC AAG ACA GTG CAT GAA GAA CTG ACC AAA TCC ATG  
ala glu pro tyr val ala ser glu tyr lys thr val his glu glu leu thr lys ser met

FIG. 4A

722 752  
 ACC ACG ATT CTC AAG CCT TCT GCC GAT TTC CTC ACC AGC AAC AAA CTA CTG AGG TAC TCA  
 thr thr ile leu lys pro ser ala asp phe leu thr ser asn lys leu leu arg tyr ser

782 812  
 TGG TTT TTC TTT GAT GTA CTG ATC AAA TCT ATG GCT CAG CAT TTG ATA GAG AAC TCC AAA  
 trp phe phe phe asp val leu ile lys ser met ala gln his leu ile glu asn ser lys

842 872  
 |Cadherin Cleavage|  
 GTT AAG TTG CTG CGA AAC CAG AGA TTT CCT GCA TCC TAT CAT CAT GCA GCG GAA ACC GTT  
 val lys leu leu arg asn gln arg phe pro ala ser tyr his his ala ala glu thr val

902 932  
 GTA AAT ATG CTG ATG CCA CAC ATC ACT CAG AAG TTT GGA GAT AAT CCA GAG GCA TCT AAG  
 val asn met leu met pro his ile thr gln lys phe gly asp asn pro glu ala ser lys

962 992  
 AAC GCG AAT CAT AGC CTT GCT GTC TTC ATC AAG AGA TGT TTC ACC TTC ATG GAC AGG GGC  
 asn ala asn his ser leu ala val phe ile lys arg cys phe thr phe met asp arg gly  
 ref 2.1 ↓

1022 1052  
 TTT GTC TTC AAG CAG ATC AAC AAC TAC ATT AGC TGT TTT GCT CCT GGA GAC CCA AAG ACC  
 phe val phe lys gln ile asn asn tyr ile ser cys phe ala pro gly asp pro lys thr

1082 1112  
 CTC TTT GAA TAC AAG TTT GAA TTT CTC CGT GTA GTG TGC AAC CAT GAA CAT TAT ATT CCG  
 leu phe glu tyr lys phe glu phe leu arg val val cys asn his glu his tyr ile pro

1142 1172  
 TTG AAC TTA CCA ATG CCA TTT GGA AAA GGC AGG ATT CAA AGA TAC CAA GAC CTC CAG CTT  
 leu asn leu pro met pro phe gly lys gly arg ile gln arg tyr gln asp leu gln leu

1202 1232  
 GAC TAC TCA TTA ACA GAT GAG TTC TGC AGA AAC CAC TTC TTG GTG GGA CTG TTA CTG AGG  
 asp tyr ser leu thr asp glu phe cys arg asn his phe leu val gly leu leu leu arg

1262 1292  
 GAG GTG GGG ACA GCC CTC CAG GAG TTC CGG GAG GTC CGT CTG ATC GCC ATC AGT GTG CTC  
 glu val gly thr ala leu gln glu phe arg glu val arg leu ile ala ile ser val leu  
 ref 3.1 ↓

1322 1352  
 AAG AAC CTG CTG ATA AAG CAT TCT TTT GAT GAC AGA TAT GCT TCA AGG AGC CAT CAG GCA  
 lys asn leu leu ile lys his ser phe asp asp arg tyr ala ser arg ser his gln ala

1382 1412/471  
 AGG ATA GCC ACC CTC TAC CTG CCT CTG TTT GGT CTG CTG ATT GAA AAC GTC CAG CGG ATC  
 arg ile ala thr leu tyr leu pro leu phe gly leu leu ile glu asn val gln arg ile

1442 1472  
 AAT GTG AGG GAT GTG TCA CCC TTC CCT GTG AAC GCG GGC ATG ACC GTG AAG GAT GAA TCC  
 asn val arg asp val ser pro phe pro val asn ala gly met thr val lys asp glu ser

1502 1532

Flt. 6A (cont.)

CTG GCT CTA CCA GCT GTG AAT CCG CTG GTG ACG CCG CAG AAG GGA AGC ACC CTG GAC AAC  
leu ala leu pro ala val asn pro leu val thr pro gln lys gly ser thr leu asp asn  
ref 4.1 and 4.2

1562 1592  
AGC CTG CAC AAG GAC CTG CTG GGC GCC ATC TCC GGC ATT GCT TCT CCA TAT ACA ACC TCA  
ser leu his lys asp leu leu gly ala ile ser gly ile ala ser pro tyr thr thr ser

1622 1652  
ACT CCA AAC ATC AAC AGT GTG AGA AAT GCT GAT TCG AGA GGA TCT CTC ATA AGC ACA GAT  
thr pro asn ile asn ser val arg asn ala asp ser arg gly ser leu ile ser thr asp  
ref 5.1 and 5.2

1682 1712  
TCG GGT AAC AGC CTT CCA GAA AGG AAT AGT GAG AAG AGC AAT TCC CTG GAT AAG CAC CAA  
ser gly asn ser leu pro glu arg asn ser glu lys ser asn ser leu asp lys his gln

1742 1772  
CAA AGT AGC ACA TTG GGA AAT TCC GTG GTT CGC TGT GAT AAA CTT GAC CAG TCT GAG ATT  
gln ser ser thr leu gly asn ser val val arg cys asp lys leu asp gln ser glu ile

1802 1832  
AAG AGC CTA CTG ATG TGT TTC CTC TAC ATC TTA AAG AGC ATG TCT GAT GAT GCT TTG TTT  
lys ser leu leu met cys phe leu tyr ile leu lys ser met ser asp asp ala leu phe

1862 1892  
ACA TAT TGG AAC AAG GCT TCA ACA TCT GAA CTT ATG GAT TTT TTT ACA ATA TCT GAA GTC  
thr tyr trp asn lys ala ser thr ser glu leu met asp phe phe thr ile ser glu val  
ref 6.1

1922 1952  
TGC CTG CAC CAG TTC CAG TAC ATG GGG AAG CGA TAC ATA GCC AGG AAC CAG GAG GGG TTG  
cys leu his gln phe gln tyr met gly lys arg tyr ile ala arg asn gln glu gly leu

1982 2012  
GGA CCC ATA GTT CAT GAT CGA AAG TCT CAG ACA TTG CCT GTT TCC CGT AAC AGA ACA GGA  
gly pro ile val his asp arg lys ser gln thr leu pro val ser arg asn arg thr gly

2042 2072  
ATG ATG CAT GCC AGA TTG CAG CAG CTG GGC AGC CTG GAT AAC TCT CTC ACT TTT AAC CAC  
met met his ala arg leu gln gln leu gly ser leu asp asn ser leu thr phe asn his

2102 2132  
AGC TAT GGC CAC TCG GAC GCA GAT GTT CTG CAC CAG TCA TTA CTT GAA GCC AAC ATT GCT  
ser tyr gly his ser asp ala asp val leu his gln ser leu leu glu ala asn ile ala

2162 2192  
ACT GAG GTT TGC CTG ACA GCT CTG GAC ACG CTT TCT CTA TTT ACA TTG GCG TTT AAG AAC  
thr glu val cys leu thr ala leu asp thr leu ser leu phe thr leu ala phe lys asn  
ref 7.1

2222 2252  
CAG CTC CTG GCC GAC CAT GGA CAT AAT CCT CTC ATG AAA AAA GTT TTT GAT GTC TAC CTG  
gln leu leu ala asp his gly his asn pro leu met lys lys val phe asp val tyr leu

2282 2312

FIG. 6A (cont.)



TGT TTT CTT CAA AAA CAT CAG TCT GAA ACG GCT TTA AAA AAT GTC TTC ACT GCC TTA AGG  
cys phe leu gln lys his gln ser glu thr ala leu lys asn val phe thr ala leu arg

2342

TCC TTA ATT TAT AAG TTT CCC TCA ACA TTC TAT GAA GGG AGA GCG GAC ATG TGT GCG GCT  
ser leu ile tyr lys phe pro ser thr phe tyr glu gly arg ala asp met cys ala ala

2372

2402

CTG TGT TAC GAG ATT CTC AAG TGC TGT AAC TCC AAG CTG AGC TCC ATC AGG ACG GAG GCC  
leu cys tyr glu ile leu lys cys cys asn ser lys leu ser ser ile arg thr glu ala

2432

2462

TCC CAG CTG CTC TAC TTC CTG ATG AGG AAC AAC TTT GAT TAC ACT GGA AAG AAG TCC TTT  
ser gln leu leu tyr phe leu met arg asn asn phe asp tyr thr gly lys lys ser phe

2492

2522

GTC CGG ACA CAT TTG CAA GTC ATC ATA TCT GTC AGC CAG CTG ATA GCA GAC GTT GTT GGC  
val arg thr his leu gln val ile ile ser val ser gln leu ile ala asp val val gly

2552

2582

ATT GGG GAA ACC AGA TTC CAG CAG TCC CTG TCC ATC ATC AAC AAC TGT GCC AAC AGT GAC  
ile gly glu thr arg phe gln gln ser leu ser ile ile asn asn cys ala asn ser asp

2612

2642

GGG CTT ATT AAG CAC ACC AGC TTC TCC TCT GAT GTG AAG GAC TTA ACC AAA AGG ATA CGC  
arg leu ile lys his thr ser phe ser ser asp val lys asp leu thr lys arg ile arg

2672

2702

ACG GTG CTA ATG GCC ACC GCC CAG ATG AAG GAG CAT GAG AAC GAC CCA GAG ATG CTG GTG  
thr val leu met ala thr ala gln met lys glu his glu asn asp pro glu met leu val

2732

2762

GAC CTC CAG TAC AGC CTG GCC AAA TCC TAT GCC AGC ACG CCC GAG CTC AGG AAG ACG TGG  
asp leu gln tyr ser leu ala lys ser tyr ala ser thr pro glu leu arg lys thr trp

2792

2822

CTC GAC AGC ATG GCC AGG ATC CAT GTC AAA AAT GGC GAT CTC TCA GAG GCA GCA ATG TGC  
leu asp ser met ala arg ile his val lys asn gly asp leu ser glu ala ala met cys

2852

xxxxxxxxxxxxxxxx Predicted

Transmembrane Domain xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxI

TAT GTC CAC GTA ACA GCC CTA GTG GCA GAA TAT CTC ACA CGG AAA GGC GTG TTT AGA CAA  
tyr val his val thr ala leu val ala glu tyr leu thr arg lys gly val phe arg gln

2942

GGG TGC ACC GCC TTC AGG GTC ATT ACC CCA AAC ATC GAC GAG GAG GCC TCC ATG ATG GAA  
gly cys thr ala phe arg val ile thr pro asn ile asp glu glu ala ser met met glu  
ref 8.1 ↓

2972

3002

GAC GTG GGG ATG CAG GAT GTC CAT TTC AAC GAG GAT GTG CTG ATG GAG CTC CTT GAG CAG  
asp val gly met gln asp val his phe asn glu asp val leu met glu leu leu glu gln

3032

3062

TGC GCA GAT GGA CTC TGG AAA GCC GAG CGC TAC GAG CTC ATC GCC GAC ATC TAC AAA CTT  
cys ala asp gly leu trp lys ala glu arg tyr glu leu ile ala asp ile tyr lys leu

3092

Fig. 6A (cont.)

ref 9.1

3122 ATC ATC CCC ATT TAT GAG AAG CGG AGG GAT TTC TTT GAA GAT GAA GAT GGA AAG GAG TAT  
ile ile pro ile tyr glu lys arg arg asp phe phe glu asp glu asp gly lys glu tyr

3152

3182 ATT TAC AAG GAA CCC AAA CTC ACA CCG CTG TCG GAA ATT TCT CAG AGA CTC CTT AAA CTG  
ile tyr lys glu pro lys leu thr pro leu ser glu ile ser gln arg leu leu lys leu

ref 10.1

3242 TAC TCG GAT AAA TTT GGT TCT GAA AAT GTC AAA ATG ATA CAG GAT TCT GGC AAG GTC AAC  
tyr ser asp lys phe gly ser glu asn val lys met ile gln asp ser gly lys val asn

3272

3302 CCT AAG GAT CTG GAT TCT AAG TAT GCA TAC ATC CAG GTG ACT CAC GTC ATC CCC TTC TTT  
pr lys asp leu asp ser lys tyr ala tyr ile gln val thr his val ile pro phe phe

3332

3362 GAC GAA AAA GAG TTG CAA GAA AGG AAA ACA GAG TTT GAG AGA TCC CAC AAC ATC CGC CGC  
asp glu lys glu leu gln glu arg lys thr glu phe glu arg ser his asn ile arg arg

3392

3422 TTC ATG TTT GAG ATG CCA TTT ACG CAG ACC GGG AAG AGG CAG GGC GGG GTG GAA GAG CAG  
phe met phe glu met pro phe thr gln thr gly lys arg gln gly gly val glu glu gln

ref 11.1

3482 TGC AAA CGG CGC ACC ATC CTG ACA GCC ATA CAC TGC TTC CCT TAT GTG AAG AAG CGC ATC  
cys lys arg arg thr ile leu thr ala ile his cys phe pro tyr val lys lys arg ile

3512

3542 CCT GTC ATG TAC CAG CAC CAC ACT GAC CTG AAC CCC ATC GAG GTG GCC ATT GAC GAG ATG  
pro val met tyr gln his his thr asp leu asn pro ile glu val ala ile asp glu met

3572 lxxxxxxx Coiled-coil 1 xxxxxx

3602 xxxxxxxx Coiled coil 1 cont'd xxxx 3632 xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx  
AGT AAG AAG GTG GCG GAG CTC CGG CAG CTG TGC TCC TCG GCC GAG GTG GAC ATG ATC AAA  
ser lys lys val ala glu leu arg gln leu cys ser ser ala glu val asp met ile lys

ref 12.1

3662 xxxxxxxxxxxxxxxxxxxxxxxx 3692 GTT CAG GTC AAT GCT GGC CCA CTA GCA TAT  
leu gln leu lys leu gln gly ser val ser val gln val asn ala gly pro leu ala tyr

3722 GCG CGA GCT TTC TTA GAT GAT ACA AAC ACA AAG CGA TAT CCT GAC AAT AAA GTG AAG CTG  
ala arg ala phe leu asp asp thr asn thr lys arg tyr pro asp asn lys val lys leu

3752

3782 CTT AAG GAA GTT TTC AGG CAA TTT GTG GAA GCT TGC GGT CAA GCC TTA GCG GTA AAC GAA  
leu lys glu val phe arg gln phe val glu ala cys gly gln ala leu ala val asn glu

lxxxxxxxxxxxxxxxxxxxxx

3842 xxxxxx Coiled coil 2 xxxxxxxxxxxx 3872 xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx  
CGT CTG ATT AAA GAA GAC CAG CTC GAG TAT CAG GAA GAA ATG AAA GCC AAC TAC AGG GAA  
arg leu ile lys glu asp gln leu glu tyr gln glu glu met lys ala asn tyr arg glu

3902 xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx 3932 xxxl

ATG GCG AAG GAG CTT TCT GAA ATC ATG CAT GAG CAG ATC TGC CCC CTG GAG GAG AAG ACG  
met ala lys glu leu ser glu ile met his glu gln ile cys pro leu glu glu lys thr

3962

AGC GTC TTA CCG AAT TCC CTT CAC ATC TTC AAC GCC ATC AGT GGG ACT CCA ACA AGC ACA  
ser val leu pro asn ser leu his ile phe asn ala ile ser gly thr pro thr ser thr

3992

4022

!xxxx PBM xxxxx!

ATG GTT CAC GGG ATG ACC AGC TCG TCT TCG GTC GTG TGA TTA CAT CTC ATG GCC CGT GTG  
met val his gly met thr ser ser ser ser val val STP

4082

TGG GGA CTT GCT TTG TCA TTT GCA AAC TCA GGA TGC TTT CCA AAG CCA ATC ACT GGG GAG

4112

4142

ACC GAG CAC AGG GAG GAC CAA GGG GAA GGG GAG AGA AAG GAA ATA AAG AAC AAC GTT ATT

4172

4202

TCT TAA CAG ACT TTC TAT AGG AGT TGT AAG AAG GTG CAC ATA TTT TTT TAA ATC TCA CTG

4232

4262

GCA ATA TTC AAA GTT TTC ATT GTG TCT TAA CAA AGG TGT GGT AGA CAC TCT TGA GCT GGA

4292

4322

CTT AGA TTT TAT TCT TCC TTG CAG AGT AGT GTT AGA ATA GAT GGC CTA CAG AAA AAA AAG

4352

4382

GTT CTG GGA TCT ACA TGG CAG GGA GGG CTG CAC TGA CAT TGA TGC CTG GGG GAC CTT TTG

4412

ref 13.1

4442

CCT CGA CTC GTG CCG GAA ATC TGA TCG TAA TCA GGG TAC AGA ACT TAC TAG TTT TGT CTA

4472

4502

GGA GTA TGT TGT ATG ACT AGG ATT TGT GCT ATT ATC TCA TTC AAC AAC ATA GAG CAA GAA

4532

4562

TAG TGA GCT AAC TGA GCT AGA CAC TCA ATT AAT CCG CTA CTG GCT TCA AGT CAG AAC TTT

4592

ref 14.1

4622

GTC ATT AAT CAT CGA CTC CGG GAC GGT CAT ATA TGT ATT ACA TTT CTA CAT TTT TAA TAC

4652

4682

TCA CAT GGG CTT ATG CAT TAA GTT TAA TTG TGA TAA ATT TGT GCT GGT CCA GTA TAT GCA

4712

4742

ATA CAC TTT AAT GGT TTA TTC TTG TCA TAA AAA TGT GCA ATA TGG AGA TGT ATA CAA GTC

4772

4802

TTT ACT

Fig. 6A (cont.)

## BAC sequences of Human CLASP 2

### Ref 1.1

Sequence of BAC4 using primer HC2AS2, which spans nucleotides 327-346 of the cDNA. Exon sequence is underlined and represents nucleotides 356-375.

TTTCTACAGNGTNTACTCAGGTATGTGCTCCTTCAACAAAATTAGCAGTTGCTGCTCTG  
TGACAAAGTTTGCACCATTTTGCAAGAAGAAAAAAATCCTAATGTGTTATATTACTATA  
TTTTTACTCTATAGATCTTTTTCTAAAGAAAGAAAGTACAACCTGAAGTGCTTATATGTA  
TTCATATAAATGACTAGTACAAGCATCATTTTGCAACAGATTTCCCCTTTCATTGGAGG  
ATCTTCTTGATGTTATTTGTACACGATCAATTTTGTAGTCTTAATAAGATGAGGCTGGGTG  
TGGTGGCTCACACCTGTAATCCTAGCATTTTGGAGGCCAAGGTGGGCAGATCACTTTAG  
CCCAGGGGTTTGAGACCAGCCTGGCCAACATGGCAAAACCTTGTCTCTACAAAAATAC  
NAAAATTATCCAGGCATGGTGATGTGTGCCTGTAGTCCCAACTNCCTAGGAGGCTAGG  
GGTAGGGGGATTGCAAGAGGGCTGGGAGGGTCAAAGCCCNAANTGAGCCATTGGTNC  
ATGTCACCTTGGACCCCAAGCNGGGGNGANCAAGAGCAAAGGACTNNTGTNNTTTAN  
AAAAAAAACCGGGCTACCATAACNACCAACCCNCNACCTACCCNACCTTTCCANNTT  
AAAANAAGGCTTTGNCTTGCANAGGAAAANCAAAATNNCC

### Ref 1.2

Sequence of BAC26 using primer HC2AS2, which spans nucleotides 327-346 of the cDNA. Exon sequence is underlined and represents nucleotides 351-375.

TCTGGTTTCTACAGTGTATACTNAGGTATGTGCTCCTTNAACAAAATTAGCAGTTGCTG  
CTCTGTGACAAAGTTTGCACCATTTTGCAAGAAGAAAAAAATCCTAATGTGTTATATTA  
CTATATTTTTACTCTATAGATCTTTTTCTAAAGAAAGAAAGTACAACCTGAAGTGCTTAT  
ATGTATTCATATAAATGACTAGTACAAGCATCATTTTGCAACAGATTTCCCCTTTCATT  
GGAGGATCTTCTTGATGTTATTTGTACACGATCAATTTTGTAGTCTTAATAAGATGAGGC  
TGGGTGTGGTGGCTCACACCTGTAATCCTAGCATTTTGGAGGCCAAGGTGGGCAGATC  
ACTTTAGCCCAGGGGTTTGAGACCAGCCTGGCCAACATGGCAAAACCTTGTCTCTACA  
AAAATACAAAAATTATCCAGGCATGGTGATGTGTGCCTGTAGTCCCAGCTACCTAGGA  
GGCTAGGGTAGGGGGATTGCAAGAGGGCTNGGAGGTCAAGGCCCGCAGTGAGCCATGG  
TCATGTCACTGCACCCCCAGCCAGGGCCGACAGGAGCAAGACTNTTGTNTCAAAAAAA  
AACAGNAACCAACANCCAACAACAACAACNACCTTTCNGCAAAANAAGCTTGCTNCA  
ANGAAACCAAAATGNCTTCTNTTTTCCCCCN

### Ref 1.3

Sequence of BAC26 using primer HC2AS2, which spans nucleotides 327-346 of the cDNA. Exon sequence is not found within this sequence. This sequence most likely represent intron sequence since this sequence matches the intron sequence found in the previous two BAC sequences.

AGNNNNNCCCNCTACNCCACTTTTAACCTTTTGAAAACACAGTGTTTNCTCAANTATGC  
GCTCCTTCACATATTAGCAGTTGCTGCTCTGTGACATAGTTGCACCATNTGCAAGAAG  
AAAAAATCCTAAGTGTNATATCACTATATNNNTACTCTATAGATCTTNTCTAAAGAAAG  
AAAGTCAACTGATGTGCTTATATGTATNCATATAAATGACTAGTACATGCATCATTTTG  
CAACAGATNTCTCCTCACATTGGAGGATCTTCTNGANGNATTCGACACGATNANTATTA  
GTCTNAATAAGATGANGCTGGTGTGGNGGTACACTGNATCTAGCATNTGGANGCATGT

Fig. 6A (cont.)

GGCAGACACTTANCCNCGGTNGAGACAGCTGTCACTGNCNAACTGTCTCTNTAAANCA  
AANNCTCCGCNNGGNGATGGGCTGAGCCAGTCCTAGNNGCTAGNTAGNGATGNNGAGN  
TGTNGCACGNCGAGNGAGCATGNTCTGTACTGACTCATCAGGCGNCNACACGNTCTGT  
TCNAAAACATACCACACACACTCNCACCTNCGCAAATTTGCTCTNNAANATGCTTNT  
TTCACACNGNTNCAATCNCTATATNNTCTTCTATTCTNACGNTNATTANNATCTTN  
CNCTGCANAACNATNCGNCCACCTNNAANNACCTTANGCTTNGTTTCACGCTTATAGCTC  
CCCTACACNTNNCAGCNNTTNCNNGTGAAGGGCCNCCCGAATCTACGANCATACTCTC  
TCCGTATATNGCCTCGGTCANCGCCATCTGCTGTNTNCTCNCNTNGCNNTTNANCNG  
TNCGCTATCTCTNNNCCGGATCCNCCNCCATATNNTNNTCTACTTANAGCGTAANNNTNT  
NCNCNCACTANTCACAACCTTNTNCTNNAACTCTATCTNCTCCTCTCTACCACCTCACT  
TACTACCTNTTCACNCANTCTCCTTCNCTNTCCACTGATCTCCACATAGCTGCTNTACTC  
GCCANTTTATCATATNCACACNCTCTACGCTNNNTNT

#### Ref 2.1

Sequence of BAC4 using primer HC2S1, which spans nucleotides 1107-1126 of the cDNA. Exon sequence is underlined and represents nucleotides 1079-1097.

CTTGTATTNAAAGAGGGTCTGCAGGAAGAAGTGTGTAGTCATAAATACCTCACTGGAT  
ATTTTATACAGGATTCTAAAAAACCTATTAGCAATAGTATGCTAGAAATAGTCATTAGC  
TTCTTGACCTTCTTAGAACTGCACACTCTATTGCACTGTACAGATTTTCAGGATGGCTGC  
AGGGATTGATTTGAAAACATAAGGACACATTTCAATAAACAATGTCTTCAATTGATTTT  
AGGGCTCCTCCTACTTCAATGAAGGACTTCAGGTAGCTTATAATTACAGACACAGGCTC  
AATACAATAAAAAAATTAGTAAGGCAGAGCTTTAAAAAAGGAAAAAGATAA  
TTCTACCAGAGAAAGGCTACATGGTGACTTCTGTTACCAGTAACAACCCCGCACTACC  
TTTGGGTCTCCAGGAGCAAAACAGCTAATGTAGTTGTTGATCTGCTTGAAGACAAAGC  
CCCTGTCCATGAAGGTGAAACATCTCTGTGGAGGAAACAAGCAAAAAAGTTATTTCA  
GGTCCAAACATTTTCGGAAATTTGGATTCAAAGCAGGCATTTATTGCTAATAAGTTTATC  
CACTGACATAAAAAACATGCCTTCAACATTGCCAGAGCACCTACTCTATTNTAGTCNCN

#### Ref 3.1

Sequence of BAC4 using primer C96AS, which spans nucleotides 1443-1452 of the cDNA. Exon sequence is underlined and represents nucleotides 1370-1422.

AATCAGCAGACCAAACAGAGGCAGGTAGAGGGTGGCTATCCTTGCCTGATGGCTCTGA  
AAAGAAGACACACATGGTAAGTTTGACCCAGGATTCTGAGAACCGAACTAAGTTGGTG  
CTGACCATCTCCTTTATTTGGATCCTTCCTATAAAGACAGATATTTGATTTTAGTCCCAA  
AATAGAGCAAAATCTTAGTGCTGTTACCATGAATTTTCTAACTGATTACTTTCTTTACAC  
CACTTAAAATAAAGGACATTATCAATGCACATTCCTTCCATTGGGGACCACTCACCCCTT  
GAAGCATATCTGTCATCAAAAGAATGCTTTATCAGCAGGTTCTTGAGCACACTGATGGC  
GATCAGACGGACCTCCCGGAACTCCTGGAGGGGCTGTCCCCACCTCCCTNAGTAACAGT  
CCCACCAAGAAGTGGTTTCTGCAGAACTCATCTGTTAATGAGTAGTCAAGCTGGGAGG  
TCTGAAATGAGGATAGAACTACTTTGNGTTAGGAAAGATGCAATGCTCTTTTGAATA  
AAACAAACAAACCAACNAACAAAAAATAAGACCCATCCTTNTGNATTTCAA  
GCCCACCCTGGGGTNGGTCAAAGAGATGATCAGNANTTTGGCNTTNAATGAAGAAAG  
AAATNAATTNTCCAGGGGNTGTTCTNCTTTTGTAGCACANGGAGGGATNTTAANTGAAA  
ACCAATTTAAATCCAATTNAGGNG

#### Ref 4.1

FIG. 6A (cont.)



Sequence of BAC4 using primer C2AS5, which spans nucleotides 1716-1735 of the cDNA. Exon sequence is underlined and represents nucleotides 1602-1703.

TTCCTTTCTGCAAGGCTGTTCCCGAATCTGTGCTTATGAGAGATCCTCTCGAATCAGCA  
TTTCTCACACTGTTGATGTTTGGAGTTGAGGTTGTATATGGAGAAGCTAAATGGAAATC  
AAGCCAACAATAAAGTTTTATTAAGACAGAACAAAATAAAGATGAGTACTGAACTTTA  
AGGGAAATTGCTTTTATTGCACTTATTTTTTCTGTTAGGAAGTTGGCTCAAGAGTTGCAT  
TCCATTACTTCACCTTTAAAGAACCAGGTCATATACAATGAGATAAAAAGAACTAGT  
CTGAAACATTGAGATGTAAACATCAATTCACCTTGTTAGAAACCACCTTTGATCGCTAAA  
GACTAAATGCATACCTGTTTCAGAATGTGATAGAATGAAGACTTAAAAAAATTAAAAG  
ATAAATCCACCTACAACCTATCAAATCACAAAATTAAACCACACAACAACTTGTAGCA  
TTCAAACCTGGTAATAAAACACTGAGGAGCCTACCCAACCTCTGAGGGGGTGTGCATGGGGTA  
TTTTAAATTTTCGAGGAGAACACAGTGATATGTGACCTCAGCCAGAAGCTGCTGTTTNA  
GCAGCAGGTTGGTGCTATGCTCCTTTTTGAAGACATATTTGTGAAGCTGGGTATTTTGG  
GGGGCCTGCTTATGATAAAAANGGCAAGGTNTTCAATGNAGGGGN

#### Ref 4.2

Sequence of BAC26 using primer C2AS5, which spans nucleotides 1716-1735 of the cDNA. Exon sequence is underlined and represents nucleotides 1602-1703.

TTCCTTTCTGGAAGGCTGTTACCCGAATCTGTGCTTATGAGAGATCCTCTCGAATCAGC  
ATTTCTCACACTGTTGATGTTTGGAGTTGAGGTTGTATATGGAGAAGCTAAATGGAAAT  
CAAGCCAACAATAAAGTTTTATTAAGACAGAACAAAATAAAGATGAGTACTGAACTTT  
AAGGGAAATTGCTTTTATTGCACTTATTTTTTCTGTTAGGAAGTTGGCTCAAGAGTTGC  
ATTCCATTACTTCACCTTTAAAGAACCAGGTCATATACAATGAGATAAAAAGAACTA  
GTCTGAAACATTGAGATGTAAACATCAATTCACCTTGTTAGAAACCACCTTTGATCGCTA  
AAGACTAAATGCATACCTGTTTCAGAATGTGATAGAATGAAGACTTAAAAAAATTAAA  
AGATAAATCCACCTACAACCTATCAAATCACAAAATTAAACCNCAACAACAACTTGTAG  
CATTCAAACCTGGTAATAAAACACTGAGGAGCCTACCCAACCTTTGAGGGGGTGTCAATGG  
GGTNTTTTTTAAATTTTTTCGNNGGGANANCCCAAGTGNTATGGTGACCTTCACCCAAGAAGC  
TTGTTTGTNTNACCAAGCNAGGTTGNNCTNTGCTCCTTTTTTAGAANACNNTATTTTNNN  
AAATNCTGGNTTTTTTNNNGNGGCCCCCTNCNTTNT

#### Ref 5.1

Sequence of BAC4 using primer C2S6, which spans nucleotides 1686-1705 of the cDNA. Exon sequence is underlined and represents nucleotides 1724-1736.

TTCCTGGATAAGGTAATTGCTTTTACCCAACACAAATGTTTCTTATAATCAATGGATTT  
AGCCCAAAGTAAACGTACTTCATGTTCTAGTGCTTTTAAAGTGTGACCTTTTGTTTTTTT  
CTAAACCACCCGGCTGACCTGGAGTAGGTGATGAGAGCTTTAAGGTTGGGGCCCATTC  
CTTGAAGTGCTCTGATTCCCTGTTTCCAGTACCTCAGATCCTGGGCAGGGTTTGCAGTGG  
AGCGTCTTGAGTGAATGGCTCTGGTGGGTTGAACGGGGAGGGACTCAAAATGCTGCCC  
ATCTCAATTTCTGTAGTCTTTTTATTATTTATTTATTTTTTGAGACAGAGTCTCGCTCT  
GTCGCCCAGGCTGGAGTACAGCGGCACGATCTCAATTNACTGCAACCTCCGCCTCC:TG  
GGTTCAAACGACTCCTCTGCCTCAGCCTCCCCAGCAGC:TGGGACCACAGGCACAAGCC  
ACCACCGCCCGGCTAATTTTTTTGTNTTTTTAGTA:GAGAT:GGGGTTTCACCATATTTGGC  
CAGGCTGGGCTCAAACCTCCTGACC:TCGTTCATCCGCNCCCTCGGNCTNCCAAAGTGCTT  
GGGATTNCAGGCNGTGAGCCCACTTACACCTNNGGCAATTCCCTGTNAGTCTTTTTTAC  
CAGAGACACCATCATTCAACACAGCTTTTCCACCCACAA

Fig. 6A (cont.)



### Ref 5.2

Sequence of BAC26 using primer C2S6, which spans nucleotides 1686-1705 of the cDNA. Exon sequence is underlined and represents nucleotides 1712-1736.

TGAGAAGAGCAATTCCTGGATAAGGTAATTGCTTTTACCCAACACAAATGTTTCTTAT  
AATCAATGGATTTAGCCCAAAGTAAACGTACTTCATGTTCTAGTGCCTTTTAAGTGTGA  
CCTTTTGTTTTTTTCTAAACCACCCGGCTGACCTGGAGTAGGTGATGAGAGCTTTAAGG  
TTGGGGCCCATTCCTTGAAGTGCTCTGATTCCTGTTTCCAGTACCTCAGATCCTGGGCA  
GGGTTTGCAGTGGAGCGTCTTGAGTGAATGGCTCTGGTGGGTTGAACGGGGAGGGACT  
CAAAATGCTGCCCCATCTCAATTCCTGTAGTCTTTTTATTTATTTATTTATTTTTTGAGAC  
AGAGTCTCGCTCTGTCGCCCAGGCTGGAGTACAGCGGCACGATCTCAATCACTGCAA  
CCTCCGNCTCCCTGGGTTCAAACGACTCCTCTGNCTNAGNCTCCC:AGCAGCCTGGGAA  
CCACAGGCTCANGCCACCACGCCCGGCTAATTNTTGTAATTTTNAGTAANAAATTGGG  
GGTTCTCACCATNTTGGCCCAAGNCTTGGGCCTAAAAACCTTNCTNACCNTCGNCATTC  
NCNCCCCNACCNTGGGCNCTNCTCAAANGNGCTTGGGGATTANCANNGGCNTTAACC  
CCCCNTATCACCGTGGNCCTTAATTT

### Ref 6.1

Sequence of BAC4 using primer C2S7, which spans nucleotides 1918-1937 of the cDNA. Exon sequence is not found within this sequence. Since the primer is directed against exon sequence we presume that sequence derived from C2S7 is intron sequence.

NAGNGNGGGTTTNAGNCGTTTGAAGCCTGNNACGNGGTGNGTGCTNGAACTCTGTGGG  
CTTTCAGGTACTGGGGTATCTGGGAGCCTGCTGTTTGCATTGCTAGTGCATCAGACCAG  
GGCTTTTTCTCCCTGTAGCTGCTACTTATACACATAGCTCTAACTGAGATGATTCTCCA  
GACAACTGATGCAGAGCAGCAAAAGCTTCTGCCGTTCTCCCCCTTCTAGGAGTGTCTCCT  
TTCTTTGGAAAGAGATCATGAGGGGCTAGATTGTAATGAAGTGAGGCTCAGTGCTTGA  
GCACATCCGGTAAAAGTTCCAATATATTGGTCATAAAGTTTCTCATTCTTTATAGCAGT  
TAATTTCTCTGGCTCATGAGTTTTCTTAGTTTTAATCTGACTTTTAAATTAATGTCTCCA  
GCACCAGTCATATCCCCAGGGCAAACCTCAAAGGCATGAGAGGCCAGACTCGGGTCCTG  
GTCATAGCAACCCCTGTCTAGGGCCTTGGTCCCTGCCTCCGCTTGTGTGCTGTGGCGCA  
GGTCCTATGGGCCCCTTAGGAAACAGGACCACCCTGTCGCACCCCCTACAGAGACCAGC  
CAAGTTTGACATTAGATCACCGTAGCAATGTNTGCAAATTCAGTTTCTTGCTAAAACA  
GGTTAAGCCTTGCAGCCACTTATCTGTAACCTGGCNGAGGTTTTGACATAAAA

### Ref 7.1

Sequence of BAC4 using primer C2S8, which spans nucleotides 2143-2162 of the cDNA. Exon sequence is underlined and represents nucleotides 2182-2219.

CTCTCGACACGCTGTTTCTATTAAACATTGGCGTTTAAGGTTTGTATCAATTTGCTGTTCTG  
NGGTTCTAGTTTTACCTTTCACATTCATTCTGCTTGGTAAGCTCAGTGAGCACAAACTTA  
CTATGTTGCATTTTTACTTCAGCAATTATTTTTGTCCCTGTAAGGAAACCATTAATCTTT  
AAATTCCTTTAATGAAATCATTCCACAGTGAATGGCTTGAATGCCCTGAAATAAAATTT  
AACTGGTCAGTGTGTGCTGCGCGCTTGGGTATGGTGGAACACGGTCTCTGGAGGCAG  
TTAACTCTTGGCTCGAACCTTGAGGATGGTGAATATAGGCACCTAATCAGGCATTTCTG  
CCTTGAATATCTTTAAATATATCCAAATGTTATAGCGTTTAATTAGATTTTTATGTAGAA  
AGGAGCAATAAACACAAGACACATGTTTTTCAGTTTTTTATCTGTTACTGCATTAAATGA

Fig. 6A (cont.)

TAAAAACGTTTTGGAGATAGAAAATGAAAGGGGGTTTTTTTTTTGTCTTGTTTTAAAGTT  
TTAGCAAATAATATTCAAGTAGGTGGAGATGGACTCTTCACCACTCTCCTGTTTTTAGG  
AACCCAATACTTTTTTCATTCTTGCTAAATGATTACTTCCATTTCTAGCATAGAAAAGGA  
GAAAATTGGAATGAGTGTTTATAT

#### Ref 8.1

Sequence of BAC4 using primer C2S9, which spans nucleotides 2992-3011 of the cDNA. Exon sequence is not found within this sequence. Since the primer is directed against exon sequence we presume that sequence derived from C2S9 is intron sequence

CGCTTTNAAATNCCAGCCGCTACTGCGGGGCGNTNAATTCGAAACGTGTTGTTNTCTGT  
GATGCCTGGCTCTGATTGTGTGGGATTGGTCATCAGTGGCGGTTGGCAGNTGGGGTTCA  
TGGAAGCGGCCATGGGGACTGATGGCAGGCCCTTGGATTGCCACCGCAGAGCCTGGCA  
GTGTCTTTGGTCTGCATTCCTACCGGCGAAGTCTCATTTACCTCACGTGTTATCTCTTG  
GAAAGCATTCCCTTTAGCGGGGCTGTGTCTACCCTTCCATCCTCTCGTCCAAACTCCCCCTC  
CTTCTCTGTTCTGTCTCCTTCCCATCCTCTTCTCCCCAGTTCTTCTTCTCTATGTTCCCTTCT  
CAGTGGTTTCTCTTCTCTGTTTGACTTTCCAAGGTCATTTTGACTGTTCCCTGCTCCCAA  
CTACAAAGATACTAAAATCTCACCTAACCCTCTTCTTCTTTCTTAATGAAAGAATGTT  
TTCAGTCCATCCCAAATTTGTGTGGACTTCACAAACCTTCTCTAAAATGGAGCCTTTTCT  
CTTCTACTCTTGACTAGNTGGTAAACGCTCCATGTTCTTGGCCAGAACTCCCTGGTGA  
GTAGCGTCACTCCCACTTTCCTGTGCAGAACCAAGCCTCCTAGAAAACCTCCTTTGCANC  
TGAGTGGGTTGGGACACGCCCTTTNTTTGGG

#### Ref 9.1

Sequence of BAC4 using primer C2AS10, which spans nucleotides 3276-3295 of the cDNA. Exon sequence is underlined and represents nucleotides 3147-3234.

TTTANACCNATNTATCCGNGTCAGTTANAGGAGTCTCTGAGAAATTTCCGACAGCGGT  
GTGAGTTTGGGTTCCCTTGTAATATACTCCTTTCCATCTTCATCTTCAAAGAATCCCTGT  
GACATAAAGCACAAATTAGAGCTATCCCTGAACGTAAGCCCAGGGCTTACCACCTAGGA  
AGCGTTCTTTTATTACAAGGGGGGAAAAAAAGGAATGGGTCTAAAAATCCAGCTGAAAT  
GGGCTTTCTGAATGAGAAAGAAAAATGCTAATAACATGAAGTCTAGGTGCAAAGGTAAA  
GGAAAAACACAACATTGCAAACCTTATTCAAGAATGCAGTCATTAAGTGTTGAGTGAAA  
TGAAAGATTTTGGATACAAGACTAAGCTGTCCCAGGGAAGTCTAATGGGAGTCAAGCC  
TGTTTCACTTTCCCAAGAAGCAGAACTCACTANAAAATGATGAGCAGCCCACGACAGG  
CAGGCTCAGAAAGTGGACATGCCTCCCTTCTCCTGATGGCTNCCATGCACACAGGATTTT  
ATGGCATGAACTGAAGCGTTTGGGGGTCTGGAGTAAGTTTAGTAAAAGTTAGGTAAAG  
CTTGTATAAATTGTATTTTGGCTTTACCCGATGAGAAAAAAAATATTNAAGACCTGGTA  
GCTTCAATATTCAAGAAAAATATTTTTCATNTCACCCG

#### Ref 10.1

Sequence of BAC4 using primer C2S11, which spans nucleotides 3167-3186 of the cDNA. Exon sequence is underlined and represents nucleotides 3231-3296.

NGNANGTGGAGCCNCGANCCAGGGACAATCTNAACCTNCTTAAACTGTACTCGGATNA  
ATTGGTTCTGAAAATGTCAAAATGATACAGGATTCTGGCAAGGTATTGACCATGTTTG  
GANAAAGTTTCATAGCAATGTAATGTTGTGATNCGATTACATATNATATATTTTTAAATG  
TNTATAGAAAAAACACANGAAAAATATTAAGGATTGTTGGCCCGTGAGTGGCAGGTG  
TATNTTCTTNCTGATCCTTTAGNGCTTTCCATTACATGCNTGACATTAAAAAAANCTTTA

Fig. 6A (cont.)

TCGCCTAATTTTTGAAACATCTAATTTTACAAAATAATTAACCGTNTGGCCANGNATAT  
TNTCATTTTTTAGGNCCAGCTATTTAGAAACTCTGACANAAATGAGGGGCTGTGGCTTNC  
CTNCCTNNACTTGNCCTCTTTTCNNGNATGTACCACATGAACTTGNCNCCTCTTTCNNC  
TNACCGGGTGGCATGTTANAGGACAGGTTGAAACNCANTNGGGCNGGANTTNGGTN  
NAATTGGGACACAATGGTACNANGCTCTATNGGAATNGAAACTCTCCCNACNNNCNGT  
GNNCCNTGGGGGAAAATGNGNCNNATTCAATTTN

#### Ref 11.1

Sequence of BAC4 using primer C2S12, which spans nucleotides 3474-3493 of the cDNA. Exon sequence is not found within this sequence. Since the primer is directed against exon sequence we presume that sequence derived from C2S9 is intron sequence

AGNANNGTTNNGCAGCTGCANNTCTGGACCCANAGGCCGCANGGGCACGAGCCNGGA  
CACGCTCGGCAAAGAGCTGTCCAGAGGGATTGAGAAGCTTCAGGACTGGAAGGGTCTT  
TCGAGCTCAGTTAGCCACCCCCACACCCATTTGAGTTTCACATTTATCTAGTGCTTCCTT  
TTGAATACTTGGGATGTTTTTCTGTTGATCTGTTGGCACTTCCTTCTTCCACAAGACCAG  
AAGCTCATATCCAATCTAAGGTCACCTTACCCTTCTGAGAATCTGATGAAAATGGCGTGC  
CTTATGTGCCTAGATGCTTTTGCACACAGTCTAAGGTGACTTATGGACTCCAGGTCCAG  
CAGCCACACCCAGTCCTGGGTCTCCGCACAGGGAGGGACCCGTCTTCACACACCTGTCT  
CAGGTTCTAGCATTGGGGCTGCTTCAGCGGTCTCAGGCTGTGAGTAAATGGGATGTGAG  
CTTGGATCGCCCCACGCTGTTGNCCCCCGGGGGGCTTGGCCAGCTGGCCACTTNGAAAT  
GCCTCCTTTTGCCCAGGAAAGCTCACTGCATTTCAATGGGGNTTNTCCACGAAGTTCAN  
CTTTANGGG

#### Ref 12.1

Sequence of BAC4 using primer C2S13, which spans nucleotides 3645-3664 of the cDNA. Exon sequence is underlined and represents nucleotides 3683-3699.

AGNAAGGTNNCTCANTNAANNCAGCGTGAGNGTTCAGGTGAGCCAGGCACAGCAGGC  
CGGAGGGCAGCAGGGGACGTCCTTGCCCCCTGGGTGACTTGAGAGTCGTTTCCACTAAC  
AAGGTCTACTTGAGAGCCTCGGTTTACCAAGTGATCCCTGCTCCCTTCCCCCAACGTNT  
GTGACATTTCTCCTGATATCAGAGGGGGAGGAAACCTCATGATCCCTGCCCCCGCCCC  
ATGAGGACTGACTGTGGGGACAAAGAGCCAGATCTCATAGACTACCCTGATTTGTCAG  
TATTTGGGGAATTCTGGGTGCCTGATTAGAAGCATCAAGACTCTTCTAAATNCAAAGA  
AGTGTGGAGAGCAGTAGATTTTCTATAAACTGGTGTTGCTGGTTTCTATGAAAATTG  
GATCCAAAAAAGTCCTTAAGTTTACCCTCTTAATGGNATCTTTTGATTAATGGAATTC  
ATTATTTTAATATAGCCCAATCAATCCAATTTTCTTTATTGGTAGCATTTTATGTTCTC  
TTTAAAAAATCTTGGNCTACCTCCAAAATTTACAGATGTTCTCCTAGGGTTTTCTCTCC  
TTTTGGTTCAAGCATCCCATTCANGTCTTGCAGTCCATTCTGGGG

#### Ref 13.1

Sequence of BAC4 using primer C2S14, which spans nucleotides 4289-4308 of the cDNA. Exon sequence is underlined and represents nucleotides 4321-4448.

GACTTANATTTATTCTTCCTTGACAGAGTAGTGTTAGAATAGATGGECTACAGAAAAAAA  
AGGTTCTGGGATCTACATGGCAGGGAGGGCTGCACTGACATTGATGCCTGGGGGACCT  
TTTGCTCGAGGCTGAGCTGGAAAATCTTGAAAATATTTTTTTTTTCTGTGGCACATTC  
AGGTTGAATACAAGAACTATTTTTGTGACTATGTTTTTGATGACCTAAGGGAACTGACC  
ATTGTAATTTTTGTACCANTGAACCANGAGATTTAAGTGCTTTTATATTCATTTCTCTTGC

Fig. 6A (cont.)

ATTTAAGAAAATATGAAAGCTTAAGGAATTATGTGAGCTTAAAACCTAGTCAAGCANTT  
TAGAACCAAAGGCCTATNTTNATAACCGCAACTATGCTNAAAAGNACAAAGTAGTACA  
GNATATTGNTATGTACATATCATTGGTAATACACNCCNGGCNTTCTGTACATATATGT  
ATTACATTTCTACNTTTTAAATACTCCCNTGGGCTTATGCCNTTAAGGTTAANTTGNGAT  
AAATTTNGGCTGTTCCNGTNTATNCNATACNCTTTT

**Ref 14.1**

Sequence of BAC4 using primer C2AS15, which spans nucleotides 4680-4700 of the cDNA. Exon sequence is underlined and represents nucleotides 4660-4683.

ATGAGAATGTAATACATATATGTACAGAATGCCAGGACTGTATTAACAATGATATGTA  
CATAACAATATACTGTACTACTTTGTACTTTTCAGCATAGTTGCGGTTATTAATATAGG  
CCTTTGGTTCTAAACTGCTTGACTAGTTTAAAGCTCACATAATTCCTTAAGCTTTCATAT  
TTTCTTAAATGCAAGGAAATGAATATAAAAGCACTAAATCTCCTGGTTCCTGGTACAA  
AAATTACAATGGTCAGTTCCCTTAGGTCATCAAAAACCTAGTCACAAAAAATAGTTCTTGT  
ATTCAACCTGAATGTGCCACAGGAAAAAAAAAATATTTTCAAGATTTTCCAGCTCAGC  
CTCGAGGCAAAAGGCCCCCAAGGCATCAATGTCAGNGCAGCCCTCCTGCCATGTAGATC  
CCAGAACCTTTTTTTTCTGTAGGCCATCTATTCTAACACTACTCTGCAGGGGAGAATAAA  
ATCTAAAGNCCAGCTCAAGAGTGCTACCAACACCTTTGTTAAGACACAATGAAAACCTT  
GGATATTGGCAGGNGAGATTTAAAAAAAAAATGTGCCCTTTCTTACCACTCCTATAGNA  
AAGTCTGGTTAAGAAATAACCGTTGGTCTTTATTTTCCTTTTNTTTCCCCTTCCCTTGGG  
NCTTCCTGGGGCTCGG

7837-101300

Fig. 6A (cont.)



HC2A -----  
KIAA ASGNLDKNARFSAIYRQDSNKLSNDDMLKLLADFRKPEKMAKLPVILGNLDITIDNVSSD  
rat -----  
HC4 -----  
HC1 -----  
HC3 -----  
HC5 -----

HC2A -----  
KIAA FPNYVNSSYIPTKQFETCSKTPITFEVEEFVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQ  
rat -----  
HC4 -----  
HC1 -----  
HC3 -----  
HC5 -----

HC2A -----VLHHHQNPEFYDEIK  
KIAA KSFAKARNIAICIEFKDSDEEDSQPLKCIYGRPGGPVFTRSAFAAVLHHHQNPEFYDEIK  
rat -----  
HC4 -----  
HC1 -----  
HC3 -----  
HC5 -----

HC2A IELPTQLHEKHHLTFFHVSCDNSSKSGSTKKRDVVETQVGYSWLPLLKDGRVVTSEQHI  
KIAA IELPTQLHEKHHLTFFHVSCDNSSKSGSTKKRDVVETQVGYSWLPLLKDGRVVTSEQHI  
rat -----  
HC4 -----  
HC1 -----  
HC3 -----  
HC5 -----

HC2A PVSANLPSGYLGYQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNNFFQYC  
KIAA PVSANLPSGYLGYQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNNFFQYC  
rat -----  
HC4 -----  
HC1 -----  
HC3 -----GPGPARSTVSISLISNSARV  
HC5 -----

HC2A OKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLT-RATQEEVAVNVTRV  
KIAA OKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLT-RATQEEVAVNVTRV  
rat -----  
HC4 -----MEIQVLIRFLSVILMQLEFWVLPNMIHEDDVPISCPMV  
HC1 -----MSFLPIILNQLFKVLV-QNEEDEITTTVTRV  
HC3 NRSRSLSNSNPDISGTPTSPDDEVRSIIGSKGLDRSNSWVNTGGPKAAPWGSNPSPSAES  
HC5 -----

Fig. 6B(cont.)

HC2A I I H V V A Q C H E S H L R S Y V K Y A Y K A E P Y V A S E Y K T V H E E L M T T I L K P S A D F L T S N  
KIAA I I H V V A Q C H E S H L R S Y V K Y A Y K A E P Y V A S E Y K T V H E E L M T T I L K P S A D F L T S N  
rat -----  
HC4 L F H I V S K C H E E G L D S Y L S S F I K Y S F R P G K P S A P Q A P L I H E T L A T M M I A L L K Q S A D F L A I N  
HC1 L P D I V A K C H E E Q L D H S V Q S Y I K F V F K T R --- A C K E R P V H E D L A K N V T G L L K - S N D S P T V K  
HC3 T Q A M D R S C N R M S S H T E T S S F L Q T L T G R L P --- T K K L F H E E L A L Q W V V C S G -- S V R --- E  
HC5 -----

Cadherin  
Cleavage

HC2A K L L R Y S W F F F D V L I K S M A Q H L I E N S K V K L I R N Q R F P A S Y H H A A E T V V N M L M P H I T Q K F G D  
KIAA K L L K Y S W F F F D V L I K S M A Q H L I E N S K V K L I R N Q R F P A S Y H H A V E T V V N M L M P H I T Q K F R D  
rat -----  
HC4 K L L K Y S W F F F E I I A K S M A T Y L L E E N K I K L T H G Q R F P K A Y H H A L H S L F L A I T - I V E S Q Y A E  
HC1 H V L K I S W F F F A I I L K S M A Q H L I D T N K I Q L E R P Q R F P E S Y Q N E L D N L V M V L S D H V I W K Y K D  
HC3 S A L Q Q A W F F F E L M V K S M V H H L Y F N D K L E A R K S R F P E R F M D D I A A L V S T I A S D I V S R F Q K  
HC5 -----

6.1  
1.1 / 1.2 / 2.1 / 2.2

HC2A N P E A S K N A N H S L A V F I K R C F T F M D R G F V F K Q I N --- N Y I S --- C F A P G D P K T L F E Y K F E F L  
KIAA N P E A S K N A N H S L A V F I K R C F T F M D R G F V F K Q I N --- N Y I S --- C F A P G D P K T L F E Y K F E F L  
rat -----  
HC4 I P K E S R N V N Y S L A S F L K C C L T L M D R G F V F N L I N --- D Y I S --- G F S P K D P K V L A E Y K F E F L  
HC1 A L E E T R R A T H S V A R F L K R C F T F M D R G C V F K M V N --- N Y I S --- M F S S G D L K T L C Q Y K F D F L  
HC3 D T E M V E R L N T S L A F F L N D L L S V M D R G F V F S L I K S C Y Q V S S K L Y S L P N P S V I V S L R L D F L  
HC5 -----

2.1  
7.1  
3.1 / 3.2

HC2A R V V C N H E H Y I P L N L P M --- P F G K G R I Q R --- Y Q D L Q L --- D Y S L T D E F  
KIAA R V V C N H E H Y I P L N L P M --- P F G K G R I Q R --- Y Q D L Q L --- D Y S L T D E F  
rat -----  
HC4 Q T I C N H E H Y I P L N L P M --- A F A K P K L Q R --- V Q D S N L --- E Y S L S D E Y  
HC1 Q E V C Q H E H F I P L C L P I R S A N I P D P L T P S E S --- T Q E L H A S D M P E Y S V T N E F  
HC3 R I I C S H E H Y V T L N L P C S L L T P P A S P S P S V S S A T S Q S S G F S T N V Q D Q K I A N M F E L S -- V P F  
HC5 ----- M N A D T A P T S P C P S I S --- S Q N S S S C S S F Q D Q K I A S M F D R T S R V P A

4.1 / 4.2

HC2A C R N H F L V G L L L R E V G T A L Q E F R E --- V R L I A I S V L K N L L I K H S F D D R Y A S R S H Q A R I A T  
KIAA C R N H F L V G L L L R E V G T A L Q E F R E --- V R L I A I S V L K N L L I K H S F D D R Y A S R S H Q A R I A T  
rat -----  
HC4 C K H H F L V G L L L R E T S I A L Q D N Y E --- I R Y T A I S V I K N L L I K H A F D T R Y Q H K N Q Q A K I A Q  
HC1 C R K H F L I G I L L R E V G F A L Q E D Q D --- V R H L A V L K N L M A K H S F D D R Y R E P R K Q A Q I A S  
HC3 R Q Q H Y L A G L V L T E L A V I L D P D A E G L F G L H K K V I N M V H N L L S S H D S D P R Y S D P Q I K A R V A M  
HC5 S S T S - S P G L L F T E L A A L D A E G E G I S E V Q R K A V S A I H S L L S S H D L D P R C V K P E V K V K I A A

3.1  
8.1

HC2A L Y L P L F G L L I E N V Q R I N V R D V S P F P V N A G - M T V K D E S L A L P A V N P L V T P Q K G S T L D N S L H  
KIAA L Y L P L F G L L I E N V Q R I N V R D V S P F P V N A G - M T V K D E S L A L P A V N P L V T P Q K G S T L D N S L H  
rat -----  
HC4 L Y L P F V G L L L E N I Q R L A G R D T L Y S C A A M P N S A S R D E F P C G --- F T S P -- A N -- R G S L S  
HC1 L Y M P L Y G M L L D N M P R I Y L K D L Y P F T V N T S N Q S R D D L S T N G G F Q S Q T A I K H A N S V D T S F S  
HC3 L Y L P L I G I I M E T V P Q L Y D F T E T H N Q R G R P I C I A T D D Y E S E --- S G --- S M I S  
HC5 L Y L P L V G I I L D A L P Q L C D F T V A D T R R Y R --- T S G S D E E Q E --- G A --- G A I T

9.1

4.1 / 4.2

HC2A K D L L G A I S G I A S P Y T T S T P N I N S V R N A D S R G S L I S T D S G N S L P E R N S E K S N S L D K H Q Q S S  
KIAA K D L L G A I S G I A S P Y T T S T P N I N S V R N A D S R G S L I S T D S G N S L P E R N S E K S N S L D K H Q Q S S  
rat -----  
HC4 T D K D T A Y G S F Q N G --- H G I K R E D S R G S L I P - E G A T G F P D Q G N T G E N --- T R Q S  
HC1 K D V L N S I A F S S --- I A I S T V N H A D S R A S L A S L D S N P S T N E K S S E K T D N C E K I P R P L  
HC3 Q T V A M A I A G T S V P Q --- L T R P G S F L L T S T S G R Q H T ---  
HC5 Q N V A L A I A G N N F N --- L K T S G - I V L S S L P Y K Q Y N ---

5.1 / 5.2  
10.1  
3.1  
2.1



HC2A	TLGNSVVRCDKLDQSEIKSLLMCFLYILKSMSDDALFTYWN-KASTSELMDFFTISEVCL	
KIAA	TLGNSVVRCDKLDQSEIKSLLMCFLYILKSMSDDALFTYWN-KASTSELMDFFTISEVCL	
rat	-----	
HC4	STRSSVSQYNRLDQYEIRSLLMCYLYIVKMISEDTLTYWN-KVSPQELINILILLEVCL	
HC1	ALIGSTLRFDRLDQAETRSLLMCFLHIMKTISYETLIAYWQ-RAPSPEVSDFFSILPVCL	11.1 / 11.2
HC3	-----TFSAESSRSLICLLWVLKN-ADETVLQKWFTDLSVLQNLRLDLLYLVCV	
HC5	-----MLNADTTRNLMICFLWIMKN-ADQSLIRKWIADLPSTQLNRILDLLFICV	
HC2A	HQFQYMGKRYIARNQEGLG--PIVHDRKS-----QTLPVSRNRTGMM	6.1
KIAA	HQFQYMGKRYIAR-----TGMM	
rat	-----	
HC4	FHFRYMGKRNIARVHDAWLSKHFGIDRKS-----QTMPALRNRSGVM	
HC1	QNFYRLGKRNIIRKIAAAF--KFVQSTQNNGTGKSNPSCQTSGLLAQWMHSTSRHEGHK	
HC3	SCFEYKGGKVFERMNSLTFK--KSKDMRAK-----LEEAILGSIGARQEMV	
HC5	LCFEYKGGQSSDKVSTQVLQ--KSRDVKAR-----LEEALLRGEGERGEMM	
HC2A	HARLQQL-----GSLDNS-----LTFNHSYGHSDADVLHQSLLEANIATEVC	
KIAA	HARLQQL-----GSLDNS-----LTFNHSYGHSDADVLHQSLLEANIATEVC	
rat	-----	
HC4	QARLQHL-----SSLESS-----FTLNHSSTTTEADIFHQALLEGNTATEVS	
HC1	QHRSQTLPIIRGK--NALSNNPKL--LQMLDNTMTSNEIDIVHHVDTEANITEGC	12.1 / 12.2
HC3	RRSRGQLERSPSGSAFGSQENLRWRKDMTHWRQNTTEKLDKSRAEIEHEALIDGNLATEAN	6.1 / 6.2
HC5	RRRAPGNDRFP-----GLNENLRWKKEQTHWRQANEKLDKTKAELDQREALISGNLATEAH	
HC2A	LTALDTLSLFTLAFKNQOLLADHGHNPLMKKVFDVYLCFLQKHQSETALKNVFTALRSLIY	7.1
KIAA	LTALDTLSLFTLAFKNQOLLADHGHNPLMKKVFDVYLCFLQKHQSETALKNVFTALRSLIY	
rat	-----KLSRGHSPLMKKVFDVYLCFLQKHQSEMAKKNVFTALRSLIY	
HC4	LTVLDTISFFTQCFKTHFLNNDGHNPLMKKVFDIHLAFLKNGQSEVSLKHVFASLRAFIS	
HC1	LTILDVLSLFTQTHQRQLQOCDCQNSLMKRGFDTYMLFFQVNSATALKHVFASLRLFVC	13.1
HC3	LIILDTLEIVVQTVS--VTES--KESILGGVLKVLHSMACNQSAVYLQHCFAQORALVS	
HC5	LIILDMQENIIQASS--ALDC--KDSLLGGVLRVLVNSLNCQSTTYLTHCFATLRALIA	3.1
HC2A	KFPSTFYEGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNDFDYTGKKSFVRTH	
KIAA	KFPSTFYEGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNDFDYTGKKSFVRTH	
rat	KFPSTFYEGRADMCASLCYEVLKCCNSKLSSIRTEASQLLYFLMRNDFDYTGKKSFVRTH	
HC4	KFPSAFFKGRVNMCAAFCEYVLKCCCTSKISSSTRNEASALLYLLMRNDFEYTKRKTFLRTH	
HC1	KFPSAFFQGPADLCGSFCYEVLKCCNHRSRSTQTEASALLYLFMRKNFEFNKQKSIVRSH	
HC3	KFPELLFEEETEQCADLCLRLLRHCSSSIGTIRSHPSASLYLLMRQNFIGN--NFARVK	7.1 / 7.2
HC5	KFGDLLFEEVEQCFDLCHQVLHHCSSSMDVTRSQCACATLYLLMRFSFGATS--NFARVK	
HC2A	LQVIISVSQLIADVVGIGETRFOQSLSIINNANCANSDRLIKHTSFSSDVKDLTRIRTVLM	
KIAA	LQVIISVSQLIADVVGIGETRFOQSLSIINNANCANSDRLIKHTSFSSDVKDLTRIRTVLM	
rat	LQVIISLSQLIADVVGIGETRFOQSLSIINNANCANSDRLIKHTSFSSDVKDLTRIRTVLM	
HC4	LQIIIAVSQLIADVALSGGSRFQESLFIINNANFANSRPMPLARAFPAEVKDLTRIRTVLM	
HC1	LQLIKAVSQLIAD-AGIGGSRFQHSALITNNFANGDKQMKNSNFPAEVKDLTRIRTVLM	14.1 / 14.2 / 15.1 / 15
HC3	MQVPMSLSSLVGTSQNFNEEFLRRSLKTIITYAEEDLELRETTFPDQVQDLVFNLMHLS	
HC5	MQVTMSLASLVGRAPDFNEEHLRRSLRTILAYSEEDTAMQMTPFPTQVEELLCNLNSILY	

HC2A	ATAQMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV	
KIAA	ATAQMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV	
rat	ATAQMKEHENDPEMLVDLQYSLAKSYASTPELRKTWLD SMARIHVKN GDLSEAAMCYVHV	
HC4	ATAQMKEHEKDPEMLIDLQYSLAKSYASTPELRKTWLD SMAKIHVKN GDFSEAAMCYVHV	
HC1	ATAQMKEHEKDPEMLVDLQYSLANSYASTPELRRTWLESMAKIHARNGDLSEAAMCYIHI	16.1 / 16.2
HC3	DTVKMKEHQEDPEMLIDLMYRIAKGYQTS PDLRLTWLQNMAGKHSERSNHAEEAQC LVHS	
HC5	DTVKMREFQEDPEMLMDL MYRIAKSYQAS PDLRLTWLQNMAEKHTKKKCYTEAMCLVHA	

		SH3	
HC2A	TALVAEYITRKGV-----		-----FRQGCTAFRVITPN
KIAA	TALVAEYITRKEA-----	VQWEPPLPHSHSACLRRSRGGVFRQGCTAFRVITPN	
rat	TALVAEYITRKEAD-----	LALQREPPVFPYSHTSCQRKSRGGMFRQGCTAFRVITPN	
HC4	AALVAEFTHRKKL-----		-----FPNGCSAFKKITPN
HC1	AALIAEYIKRKG YWKVEKICTASLLSE DTHPCDSNSLLTTPSGGSMFSMGWPAFLSITPN		
HC3	AALVAEYISMLED-----		-----RKYLPVGCVTFNIS SN
HC5	AALVAEYISMLED-----		-----HSYLPVGSVSFQNI SSN

HC2A	IDEEASMMEDVGMQD-----	VHFNEDVLMELLEQCADGLWKAERYELIADIYKLIIP I	9.1
KIAA	IDEEASMMEDVGMQD-----	VHFNEDVLMELLEQCADGLWKAERYELIADIYKLIIP I	
rat	IDEEASMMEDVGMQD-----	VHFNEDVLMELLEQCADGLWKAERYELIADIYKLIIP I	
HC4	IDEEGAMKEDAGMMD-----	VHYSEEV LLELLEQCVNGLWKAERYEI ISEISKLIGPI	
HC1	IKEEGA AKEDSGMHD-----	TPYNE NILVEQLYMCGEFLWK SERYELIADV NKP IIAV	17.1 / 17.2
HC3	VLEESAVSDDV VSPDEEGICSGKYFTESGLVGLLEQAAASF SMAGMYEAVNEVYKVLIP I		
HC5	VLEESV VSED T LSPDE DGVCAGQYFTESGLVGLLEQAAELFSTGGLYETVNEVYKLVIP I		

		ITAM	ITAM	ITAM	ITAM	
HC2A	YEKRRD-----					9.1
KIAA	YEKRRD FERLAHLYDTLH RAYSKVTEVMHSGRLLGTYFRVAFFGQAAQYQFTDSETDVE					
rat	SMKSGGTLETTHLYDTLH REYSKVTEVITR-----			A-----	AGSWDLLPGGLFGQ	
HC4	YENRREFENLTQVYRTLHGAYTKILEVMHTKKRLLG-----				TFERVAFYGO	
HC1	FEKQRDFKKLS DLYDIHRSYLKVAE VVNSEKRLFG-----				RYRVAFYGO	
HC3	HEANRDAKKLSTIHGKLQEA FSKIVHQSTGWERMFG-----				TYFRVGFYGO	9.1
HC5	LEAHREFRKLTLTHSKLQRA FDSIVNKDH--KRMFG-----				TYFRVGFYGO	

HC2A	-FFED EDGKEYIYKEPKLTP LSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYA	10.1
KIAA	GFFED EDGKEYIYKEPKLTP LSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKYA	
rat	GFFED EDGKEYIYKEPKLTP LSEISQRLLKLYSDKFGSENVKMIQDSGKVNPKDLDSKFA	
HC4	SFFEE EDGKEYIYKEPKLTGLSEISLRLVKLYGEKFGTENVKI IQDSDKVNAKELDPKYA	
HC1	GFFEE EEGKEYIYKEPKLTGLSEISQRLLKLYADKFGADNVKI IQDSNKNPNKDLDPKYA	
HC3	TKFGDLDEQEFVYKEPAITKLAEI SHRLEGFYGERFGEDVVEVIKDSNPFVDKCKLDPNKA	10.1 / 10.2
HC5	SKFGDLDEQEFVYKEPAITKLPEI SHRLEAFYGC FGAEFVEVIK DSTPVDKTKLDPNKA	4.1

HC2A	YIQVTHVIPFFDEKELQERKTEFERSHNIRRFMFEMPFTQTGKRQGGVVEEQCKRRTILTA	11.1 / 11.2
KIAA	YIQVTHVIPFFDEKELQERKTEFERSHNIRRFMFEMPFTQTGKRQGGVVEEQCKRRTILTA	
rat	YIQVTHVTPFFDEKELQERKTEFERCHNIRRFMFEMPFTQTGKRQGGVVEEQCKRRTILTA	
HC4	HIQVTYVKPYFDDKELTERKTEFERNHNISR FVFEAPYTL SGKKQGCIEEQCKRRTILT	
HC1	YIQVTYVTPFFEEKEIEDRKTDFEMHNNIRRFVFETPFTLSGKKHGGVAEQCKRRTILT	18.1
HC3	YIQITYVEPYFDTYEMKDRITYFDKNYNLRRFMYCTPFTLDGRAHGE LHEQFKRRTILT	
HC5	YIQITFVEPYFDEYEMKDRVITYFEKNFNLRRFMYTTPFTLEGRPRGELHEQYRRNTVLT	

Coiled-Coil

HC2A	IHC	FPYVKKR	IPV	MYQHHTDLN	IEVAIDEMSKKVAELRQLCSSAEVDMIKLQKLQGSV
KIAA	IHC	FPYVKKR	IPV	MYQHHTDLN	IEVAIDEMSKKVAELRQLCSSAEVDMIKLQKLQGSV
rat	IHC	FPYVKKR	IPV	MYQHHTDLN	IEVAIDEMSKKVAELHQLCSSAEVDMIKLQKLQGSV
HC4	SNS	FPYVKKR	IPIN	CEQQINLKE	IDGATDEIKDKTAELOKLCSSSTDVDMIQLQKLQGSV
HC1	SHL	FPYVKKR	IQVIS	QSSTELN	IEVAIDEMSRKVSELNQLCTMEEVDMISLQKLQGSV
HC3	SHAF	PIKTRVNV	THKEE	IIILTH	IEVAIEDMQKKTQELAFATHQDPADPKMLQMLQGSV
HC5	MHAF	PIKTRISV	IQKEEF	VLTH	IEVAIEDMKKKTLOLAVAINQEPDAKMLQMLQGSV

11.1

Coiled-Coil 2

HC2A	SVQ	VNAGPLAYARAFLDDTNTKRYPDNKVKLLKEVFRQFVEACGOALAVNERLIKEDQLE
KIAA	SVQ	VNAGPLAYARAFLDDTNTKRYPDNKVKLLKEVFRQFVEACGOALAVNERLIKEDQLE
rat	SVQ	VNAGPLAYARAFLDDTNTKRYPDNKVKLLKEVFRQFVEACGOALAVNERLIKEDQLE
HC4	SVQ	VNAGPLAYARAFLNDSQASKYPPKKVSELKDMFRKFQACSI
HC1	SVK	VNAGPMAYARAFLEETNAKKYPDNQVKLLKEIFRQFADACGOALDVNERLIKEDQLE
HC3	GTT	VNOGPLEVAQVFLSEIPSDPKLFRHHNKLRLCFKDFTKRCEDALRKNKSLIGPVQKE
HC5	GAT	VNOGPLEVAQVFLAEIPADPKLYRHHNKLRLCFKEFIMRCGEAVEKNKRLITADQRE

11.1 / 12.1

Coiled-Coil 2

HC2A	YQE	EMKANYREMAKELSEIMHEQICPLEEKTS-VLPNSLHIFNAISGTPTSTMVHGMTSS
KIAA	YQE	EMKANYREMAKELSEIMHEQLG-----
rat	YQE	EMKANYREIRKELSDIIVPRICPGEDKRATKFPAPHLQRHQRTDNKHSGSRVDQFELS
HC4	YHE	GLKSNFRDMVKELSDIIEHQILOEDTMHSPWMSNTLHVFCASGTSSDRGYGSPRYA
HC1	YQE	ELRSHYKDMSELSTVMNEQITGRDDLK---RGVDQCTRVISKATPALPTVSISS
HC3	YQRE	LG----KLSS-----PZ-----
HC5	YQEL	KKNYNKLKENLRPMIERKIPELYKPIFRVESQKRDSFHRSSFRCETQLSQGSZ-

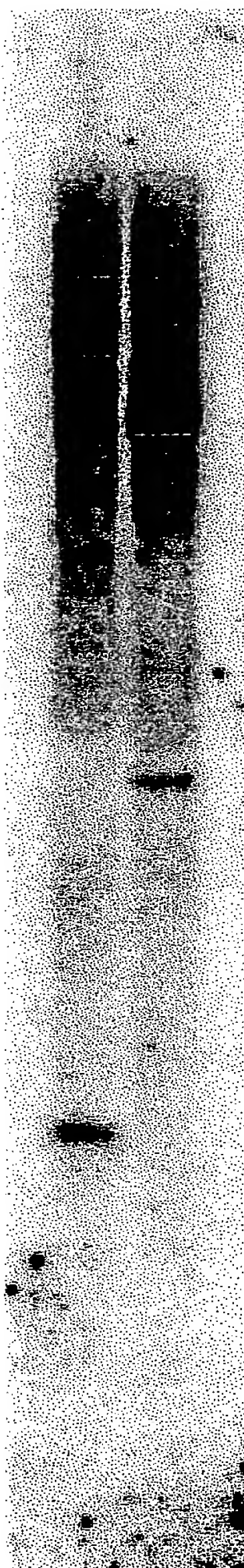
19.1

PBM	SSVV	-----
HC2A	CVT	LPHEPHVGTCTFVMCKLRTTFRANHWFCQAQEEAMGNGREKEPWTVIFNSRFYRSWKG
KIAA	EV	-----
rat	SAEV	-----
HC4		-----
HC1		-----
HC3		-----
HC5		-----

HC2A	-----
KIAA	-----
rat	VHIFF
HC4	-----
HC1	-----
HC3	-----
HC5	-----

096833-101900

HinDIII  
Eco RI

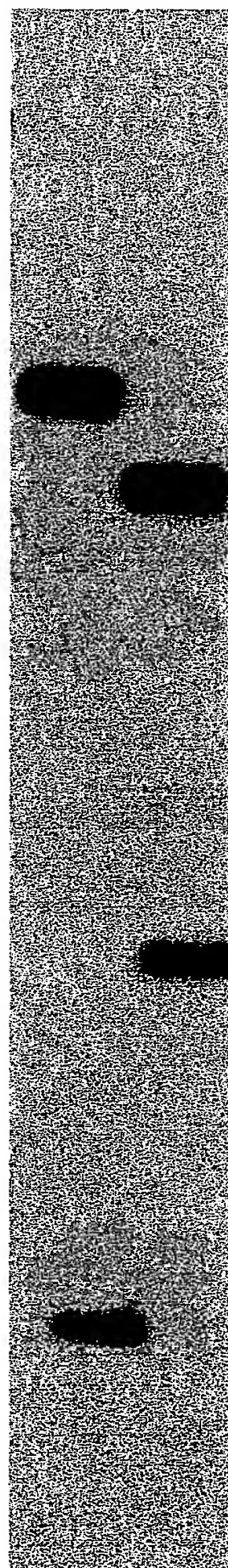


← ~ 4.5 kb

← ~ 1.85 kb

genomic DNA

Pst I  
Eco RI



← ~ 4.5 kb

← ~ 1.85 kb

BAC 6 DNA

Fig. 7



09687837-101300

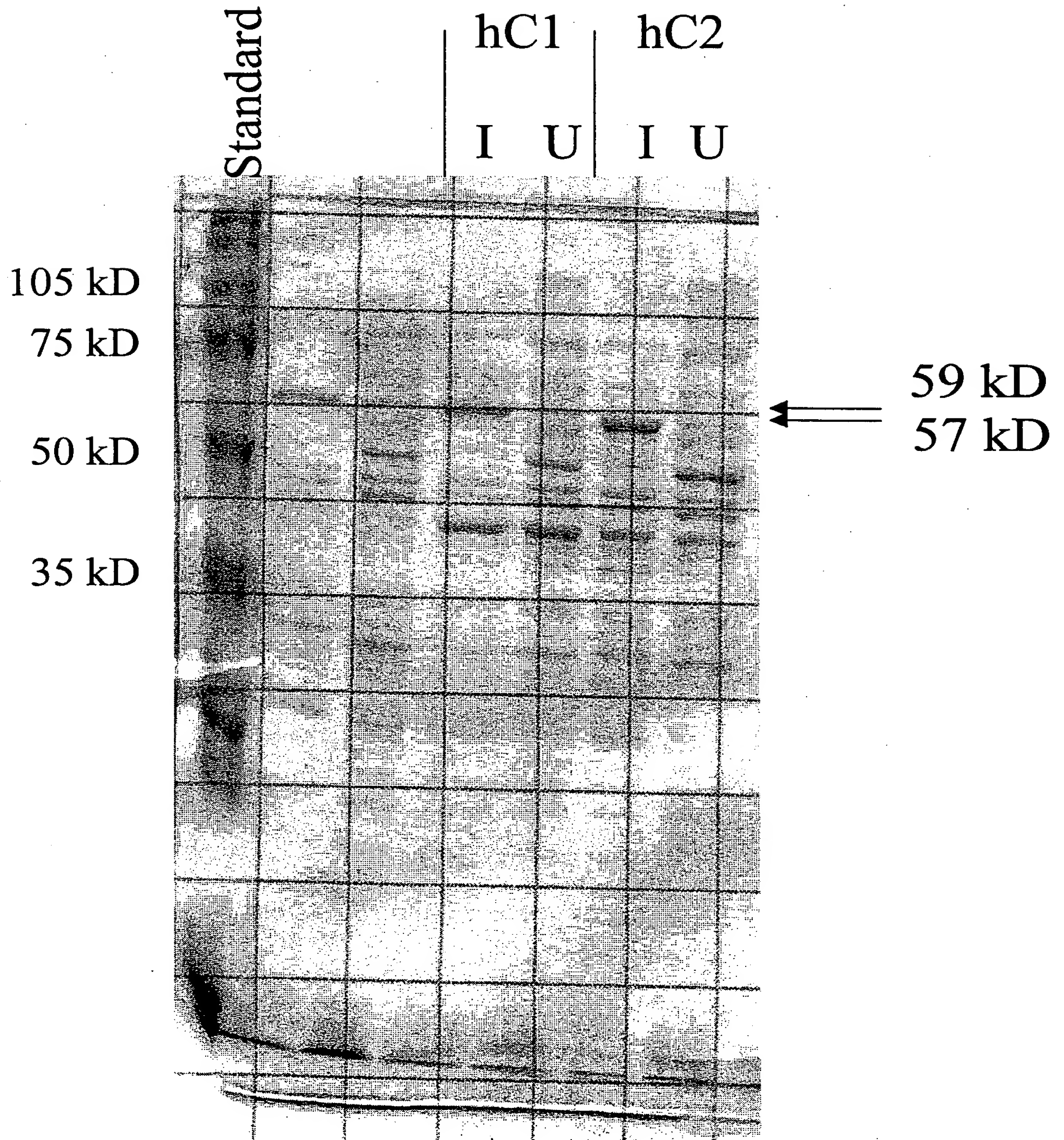


FIG. 8

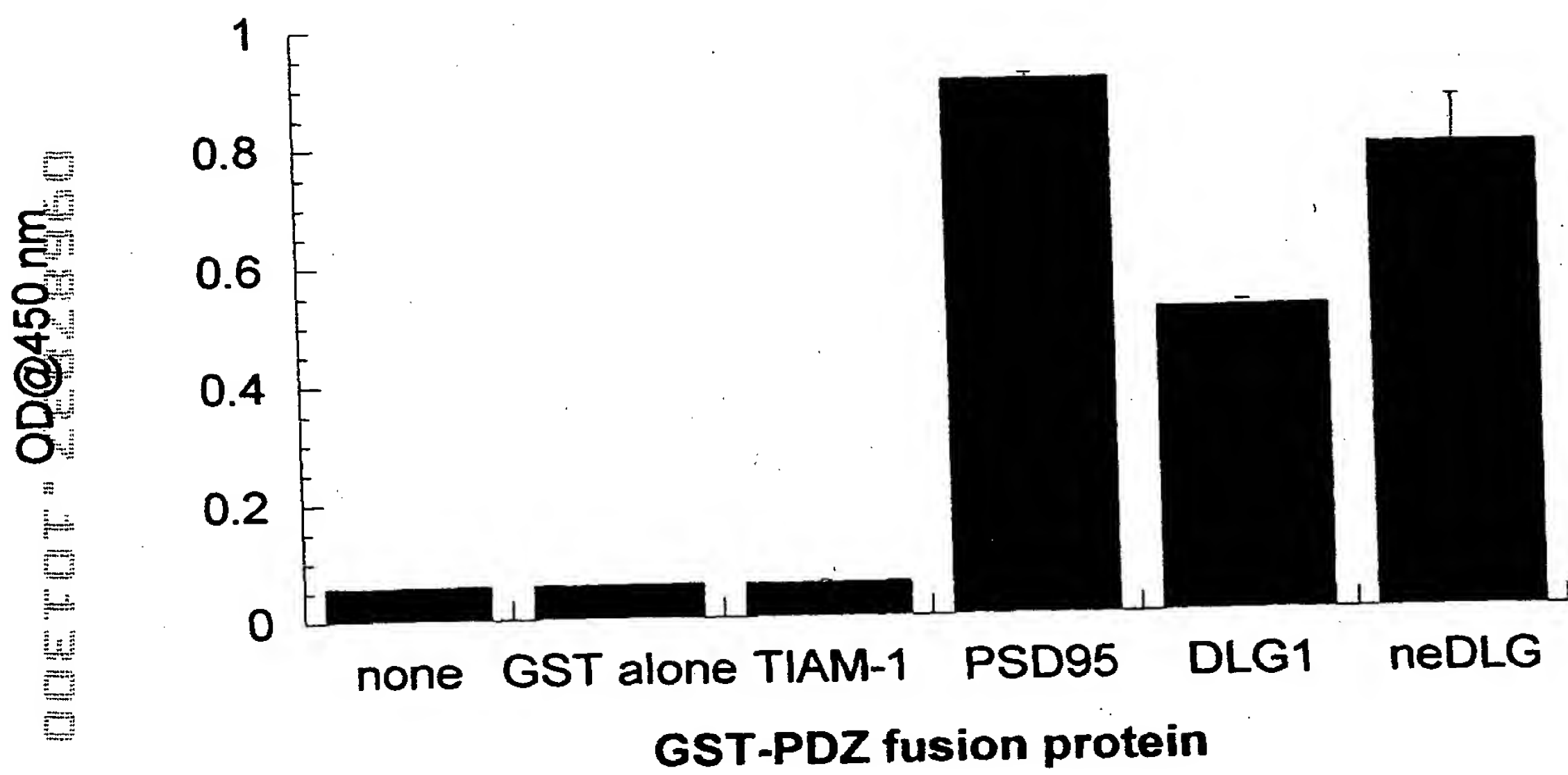


Fig. 9A



000000-22829864  
fraction maximal binding

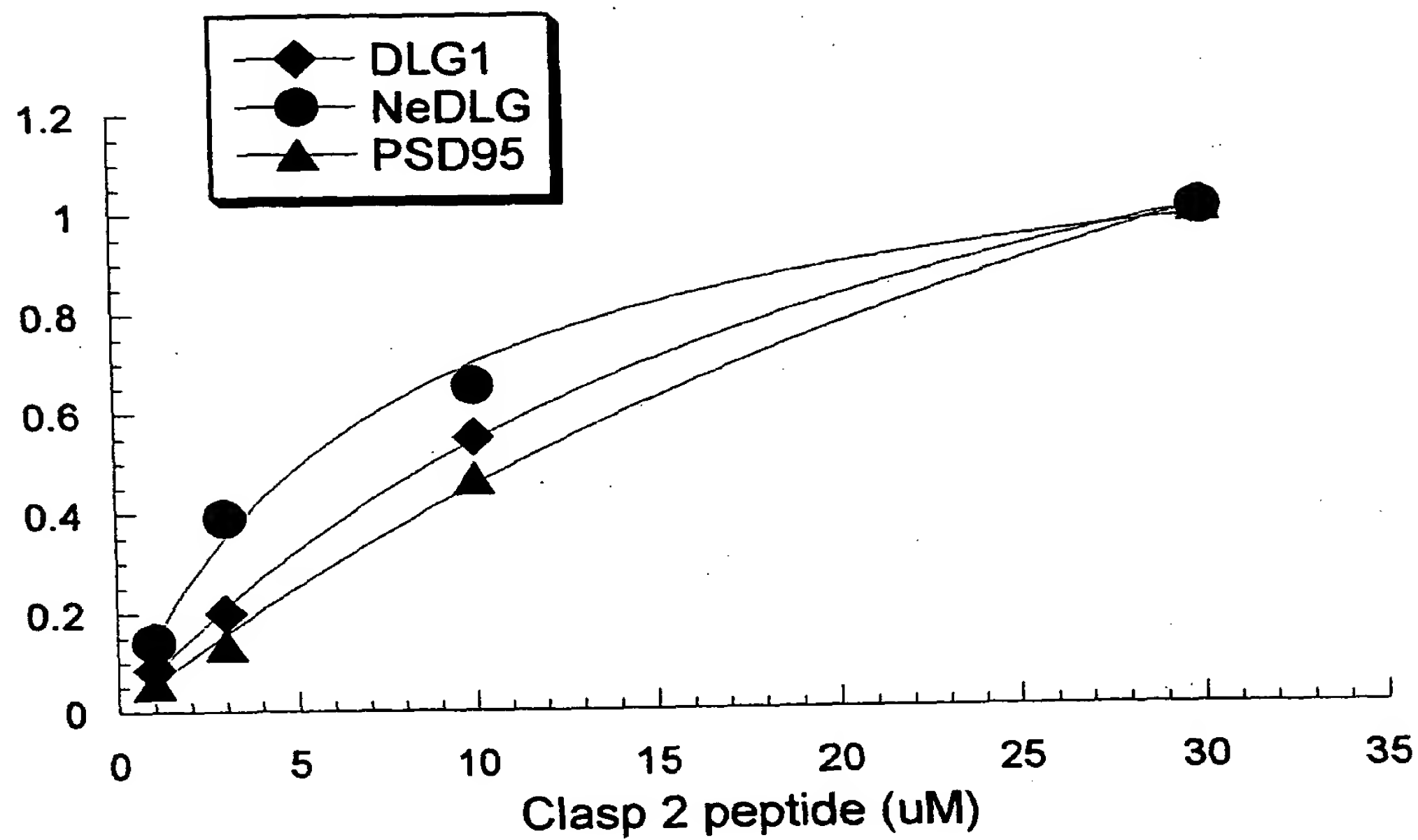


FIG. 9B

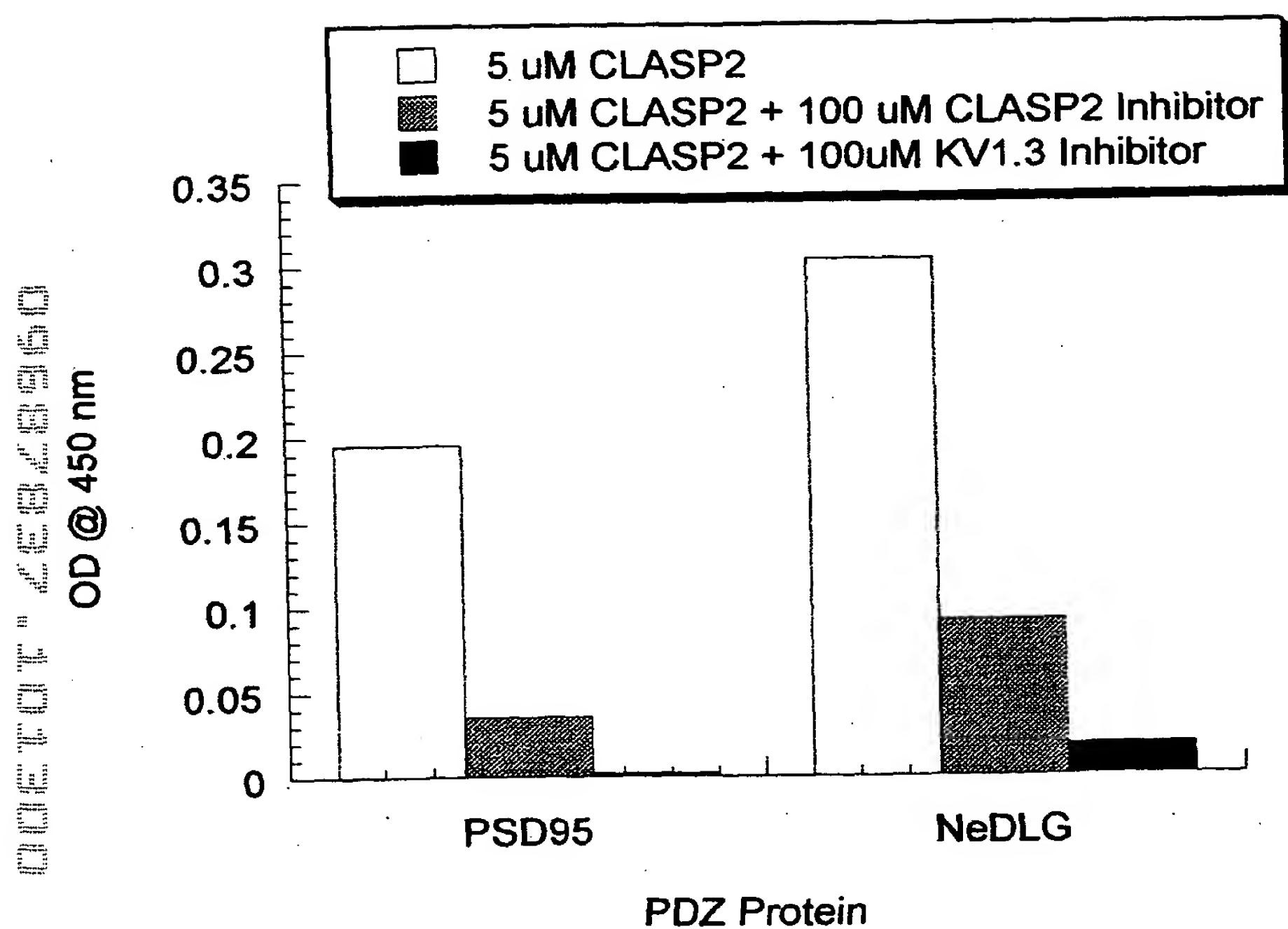


Fig. 9C

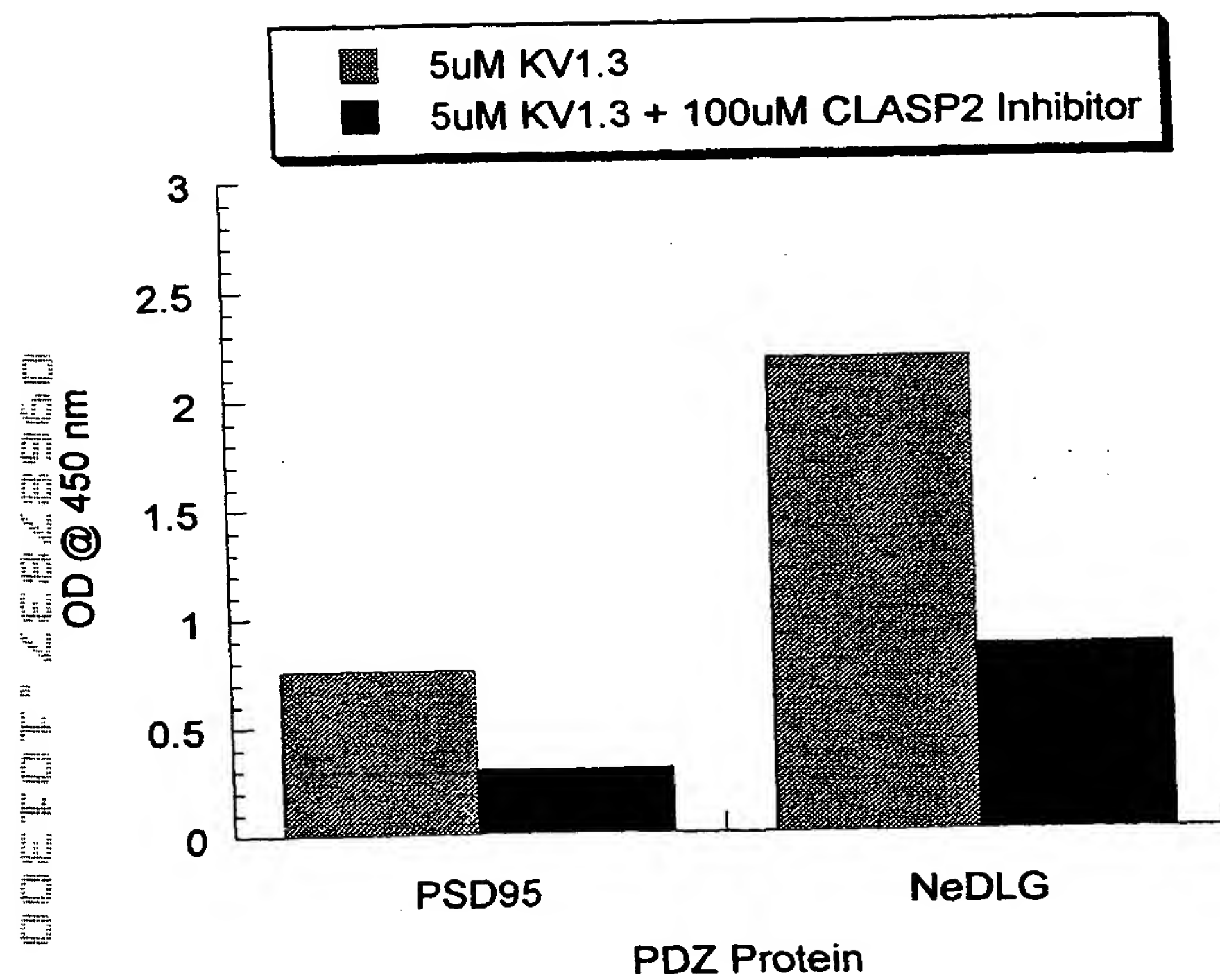


FIG. 9D

	10	20	30	40	50	60	70	80
1	AATTGTAATA	CGACTCACTA	TAGGGCGAAT	TGGGTACCGG	CCCCCCCCTC	GAGGTGACAG	GTATCGATAA	GCTTGATATC 80
81	GAATTCGGCA	CGAGTTTTAC	ACCATCACCA	AAACCCAGAA	TTTTATGATG	AGATTAAAAT	AGAGTTGCCC	ACTCAGCTGC 160
161	ATGAAAAGCA	CCACCTGTTG	CTCACATTCT	TCCATGTCAG	CTGTGACAAC	TCAAGTAAAG	GAAGCACGAA	GAAGAGGGAT 240
241	GTCGTTGAAA	CCCAAGTTGG	CTACTCCTGG	CTTCCCCTCC	TGAAAGACGG	AAGGGTGGTG	ACAAGCGAGC	AGCACATCCC 320
321	GGTCTCGGCG	AACCTTCCTT	CGGGCTATCT	TGGCTACCAA	GAGCTTGGGA	TGGGCAGGCA	TTATGGTCCG	GAAATTAAAT 400
401	GGGTAGATGG	AGGCAAGCCA	CTGCTGAAAA	TTTCCACTCA	TCTGGTTTCT	ACAGGGATAC	TCAGGATCAG	CATTTACATA 480
481	ATTTTTTCCA	GTACTGTCAG	AAAACCGAAT	CTGGAGCCCA	AGCCTTAGGA	AACGAACTTG	TAAAGTACCT	TAAGAGTCTG 560
561	CATGCGATGG	AAGGCCACGT	GATGATCGCC	TTCTTGCCCA	CTATCCTAAA	CCAGCTGTTC	CGAGTCTCTA	CCAGAGCCAC 640
641	ACAGGAAGAA	GTCGCGGTTA	ACGTGACTCG	GGTCATTATT	CATGTGGTTG	CCCAGTGCCA	TGAGGAAGGA	TTGGAGAGCC 720
721	ACTTGAGGTC	ATATGTTAAG	TACGCGTATA	AGGCTGAGCC	ATATGTTGCC	TCTGAATACA	AGACAGTGCA	TGAAGAAGTG 800
801	ACCAAATCCA	TGACCACGAT	TCTCAAGCCT	TCTGCCGATT	TCCTCACCAG	CAACAACTA	CTGAGGTACT	CATGGTTTTT 880
881	CTTTGATGTA	CTGATCAAAT	CTATGGCTCA	GCATTTGATA	GAGAACTCCA	AAGTTAAGTT	GCTGCGAAAC	CAGAGATTTT 960
961	CTGCATCCTA	TCATCATGCA	GCGGAAACCG	TTGTAAATAT	GCTGATGCCA	CACATCACTC	AGAAGTTTGG	AGATAATCCA 1040
1041	GAGGCATCTA	AGAAGCGGAA	TCATAGCCTT	GCTGTCTTCA	TCAAGAGATG	TTTCACCTTC	ATGGACAGGG	GCTTTGTCCT 1120
1121	CAAGCAGATC	AACAACCTACA	TTAGCTGTTT	TGCTCCTGGA	GACCCAAAGA	CCCTCTTTGA	ATACAAGTTT	GAATTTCTCC 1200
1201	GTGTAGTGTG	CAACCATGAA	CATTATATTC	CGTTGAACTT	ACCAATGCCA	TTTGGAAGAG	GCAGGATTCA	AAGATACCAA 1280
1281	GACCTCCAGC	TTGACTACTC	ATTAACAGAT	GAGTTCTGCA	GAAACCACTT	CTTGGTGGGA	CTGTTACTGA	GGGAGGTGGG 1360
1361	GACAGCCCTC	CAGGAGTTCC	GGGAGGTCCG	TCTGATCGCC	ATCAGTGTGC	TCAAGAACCT	GCTGATAAAG	CATTCTTTTG 1440
1441	ATGACAGATA	TGCTTCAAGG	AGCCATCAGG	CAAGGATAGC	CACCTCTAC	CTGCCTCTGT	TTGGTCTGCT	GATTGAAAAC 1520
1521	GTCCAGCGGA	TCAATGTGAG	GGATGTGTCA	CCCTTCCCTG	TGAACGCGGG	CATGACCGTG	AAGGATGAAT	CCCTGGCTCT 1600
1601	ACCAGCTGTG	AATCCGCTGG	TGACGCCGCA	GAAGGGAAGC	ACCCTGGACA	ACAGCCTGCA	CAAGGACCTG	CTGGGCGCCA 1680
1681	TCTCCGGCAT	TGCTTCTCCA	TATACAACCT	CAACTCCAAA	CATCAACAGT	GTGAGAAATG	CTGATTCCAG	AGGATCTCTC 1760
1761	ATAAGCACAG	ATTCCGGTAA	CAGCCTTCCA	GAAAGGAATA	GTGAGAAGAG	CAATTCCCTG	GATAAGCACC	AACAAAGTAG 1840
1841	CACATTGGGA	AATTCCGTGG	TTCCGTGTGA	TAAACTTGAC	CAGTCTGAGA	TTAAGAGCCT	ACTGATGTGT	TTCTCTTACA 1920
1921	TCITAAAGAG	CATGTCTGAT	GATGCTTTGT	TTACATATTG	GAACAAGGCT	TCAACATCTG	AACTTATGGA	TTTTTTTACA 2000
2001	ATATCTGAAG	TCGCTCTGCA	CCAGTTCCAG	TACATGGGGA	AGCGATACAT	AGCCAGGAAC	CAGGAGGGGT	TGGGACCCAT 2080
2081	AGTTCATGAT	CGAAAGTCTC	AGACATTGCC	TGTTTCCCGT	AACAGAACAG	GAATGATGCA	TGCCAGATTG	CAGCAGCTGG 2160
2161	GCAGCCTGGA	TAACCTCTCT	ACTTTTAACC	ACAGCTATGG	CCACTCGGAC	GCAGATGTTT	TGCACCAGTC	ATTACTTGAA 2240
2241	GCCAAACATT	CTACTGAGGT	TTGCCTGACA	GCTCTGGACA	CGCTTTCTCT	ATTTACATTG	GCGTTTAAGA	ACCAGCTCCT 2320
2321	GGCCGACCAT	GGACATAATC	CTCTCATGAA	AAAAGTTTTT	GATGTCTACC	TGTGTTTTCT	TCAAAAACAT	CAGTCTGAAA 2400
2401	CGGCTTTAAA	AAATGTCTTC	ACTGCCTTAA	GGTCCCTTAAT	TTATAAGTTT	CCCTCAACAT	TCTATGAAGG	GAGAGCGGAC 2480
2481	ATGTGTGCGG	CTCTGTGTTA	CGAGATTCTC	AAGTGTGTGA	ACTCCAAGCT	GAGCTCCATC	AGGACGGAGG	CCTCCCAGCT 2560
2561	GCTCTACTTC	CTGATGAGGA	ACAACCTTGA	TTACACTGGA	AAGAAGTCCT	TTGTCCGGAC	ACATTTGCAA	GTCATCATAT 2640
2641	CTGTACGCCA	GCTGATGACA	GACGTTGTTG	GCATTGGGGA	AACCAGATTC	CAGCAGTCCC	TGTCCATCAT	CAACAACGTG 2720
2721	GCCAAACAGT	ACCGGCTTAT	TAAGCACACC	AGCTTCTCCT	CTGATGTGAA	GGACTTAACC	AAAAGGATAC	GCACGGTGCT 2800
2801	AATGGCCACC	GCCCAGATGA	AGGAGCATGA	GAACGACCCA	GAGATGCTGG	TGGACCTCCA	GTACAGCCTG	GCCAAATCCT 2880
2881	ATGCCAGCAC	GCCCAGAGCT	AGGAAGACGT	GGCTCGACAG	CATGGCCAGG	ATCCATGTCA	AAAATGGCGA	TCTCTCAGAG 2960
2961	GCAGCAATGT	GCTATGTCCA	CGTAACAGCC	CTAGTGGCAG	AATATCTCAC	ACGGAAGGCG	GTGTTTAGAC	AAGGATGCAC 3040
3041	CGCCTTCAGG	GTCATTACCC	CAAACATCGA	CGAGGAGGCC	TCCATGATGG	AAGACGTGGG	GATGCAGGAT	GTCCATTTCA 3120
3121	ACGAGGATGT	GCTGATGGAG	CTCCTTGAGC	AGTGCGCAGA	TGGACTCTGG	AAAGCCGAGC	GCTACGAGCT	CATCGCCGAC 3200
3201	ATCTACAAAC	TTATCATCCC	CATTTATGAG	AAGCGGAGGG	ATTTCTTTGA	AGATGAAGAT	GGAAAGGAGT	ATATTTACAA 3280
3281	GGAACCCAAA	CTCACACCGC	TGTCGGAAAT	TTCTCAGAGA	CTCCTTAAAC	TGTACTCGGA	TAAATTTGGT	TCTGAAAATG 3360
3361	TCAAAATGAT	ACAGGATTCT	GGCAAGGTCA	ACCCTAAGGA	TCTGGATTCT	AAGTATGCAT	ACATCCAGGT	GACTCACGTC 3440
3441	ATCCCTTTCT	TTGACGAAAA	AGAGTTGCAA	GAAAGGAAAA	CAGAGTTTGA	GAGATCCCAC	AACATCCGCC	GCTTCATGTT 3520
3521	TGAGATGCCA	TTTACGCAGA	CCGGGAAGAG	GCAAGGCGGG	GTGGAAGAGC	AGTGCAAACG	GCGCACCATC	CTGACAGCCA 3600
3601	TACACTGCTT	CCCTTATGTG	AAGAAGCGCA	TCCCTGTCTAT	GTGCTCCTC	GGCCGAGGTG	GACATGATCA	AACTGCAGCT 3760
3681	ATTGACGAGA	TGAGTAAGAA	GGTGGCGGAG	CTCCGGCAGC	CCACTAGCAT	ATGCGCGAGC	TTTCTTAGAT	GATACAAACA 3840
3761	CAAACCTCCAG	GGCAGCGTGA	GTGTTCAAGT	CAATGCTGGC	AGTTTTTCAGG	CAATTTGTGG	AAGCTTGCGG	TCAAGCCTTA 3920
3841	CAAAGCGATA	TCCTGACAAT	AAAGTGAAGC	TGCTTAAGGA	ATCAGGAAGA	AATGAAAGCC	AACTACAGGG	AAATGGCGAA 4000
3921	GCGGTAAACG	AACGTCTGAT	TAAAGAAGAC	CAGCTCGAGT	GAGGAGAAGA	CGAGCGTCTT	ACCGAATTCC	CTTCACATCT 4080
4001	GGAGCTTTCT	GAAATCATGC	ATGAGCAGAT	CTGCCCCCTG	CGGGATGACC	AGCTCGTCTT	CGGTCTGTGT	ATTACATCTC 4160
4081	TCAACGCCAT	CAGTGGGACT	CCAACAAGCA	CAATGGTTCA	CAGGATGCTT	TCCAAAGCCA	ATCACTGGGG	AGACCGAGCA 4240
4161	ATGGCCCGTG	TGTGGGGACT	TGCTTTGTCA	TTTGCAAACT	AACAACGTTA	TTTCTTAACA	GACTTTCTAT	AGGAGTTGTA 4320
4241	CAGGGAGGAC	CAAGGGGAAG	GGGAGAGAAA	GGAAATAAAG	CAAAGTTTTT	ATTGTGTCTT	AACAAAGGTG	TGGTAGACAC 4400
4321	AGAAGGTGCA	CATATTTTTT	TAAATCTCAC	TGGCAATATT	GTGTTAGAAT	AGATGGCCTA	CAGAAAAAAA	AGGTTCTGGG 4480
4401	TCTTGAGCTG	GACTTAGATT	TTATTCTTCC	TTGCAGAGTA	GGGGACCTTT	TGCCTCGACT	CGTGCCGGAA	ATCTGATCGT 4560
4481	ATCTACATGG	CAGGGAGGGC	TGCACTGACA	TTGATGCCCTG	TTGTATGACT	AGGATTTGTG	CTATTATCTC	ATTCAACAAC 4640
4561	AATCAGGGTA	CAGAACTTAC	TAGTTTTGTC	TAGGAGTATG	TTAATCCGCT	ACTGGCTTCA	AGTCAGAACT	TTGTCAATTA 4720
4641	ATAGAGCAAG	AATAGTGAGC	TAAGTGAGCT	AGACACTCAA	CATTTTTAAT	ACTCACATGG	GCTTATGCAT	TAAGTTTAAAT 4800
4721	TCATCGACTC	CGGGACGGTC	ATATATGTAT	TACATTTCTA	TAATGGTTTA	TTCTTGTCAT	AAAAATGTGC	AATATGGAGA 4880
4801	TGTGATAAAT	TTGTGCTGGT	CCAGTATATG	CAATACACTT				
4881	TGTATACAAG	TCTTTACT						4898

FIG. 10A



	10	20	30	40	50	60	70	80
1	AATTGTAATA	CGACTCACTA	TAGGGCGAAT	TGGGTACCGG	GCCCCCCTC	GAGGTCGACG	GTATCGATAA	GCTTGATATC
81	GAATTCGGCA	CGAGTTTTAC	ACCATCACCA	AAACCCAGAA	TTTTATGATG	AGATTAAAT	AGAGTTGCC	ACTCAGCTGC
161	ATGAAAAGCA	CCACCTGTTG	CTCACATTCT	TCCATGTCAG	CTGTGACAAC	TCAAGTAAAG	GAAGCAGGAA	GAAGAGGGAT
241	GTCGTTGAAA	CCCAAGTTGG	CTACTCCTGG	CTTCCCCTCC	TGAAAGACGG	AAGGGTGGTG	ACAAGCGAGC	AGCACATCCC
321	GGTCTCGGCG	AACCTTCCTT	CGGGCTATCT	TGGCTACCAA	GAGCTTGGGA	TGGGCAGGCA	TTATGGTCCG	GAAATTAAAT
401	GGGTAGATGG	AGGCAAGCCA	CTGCTGAAAA	TTTCCACTCA	TCTGGTTTCT	ACAGGGATAC	TCAGGATCAG	CATTTACATA
481	ATTTTTTCCA	GTACTGTCAG	AAAACCGAAT	CTGGAGCCCA	AGCCTTAGGA	AACGAACCTG	TAAAGTACCT	TAAGAGTCTG
561	CATGCGATGG	AAGGCCACGT	GATGATCGCC	TTCTTGCCCA	CTATCCTAAA	CCAGCTGTTT	CGAGTCCTCA	CCAGAGCCAC
641	ACAGGAAGAA	GTCGCGGTTA	ACGTGACTCG	GGTCATTATT	CATGTGGTTG	CCCAGTGCCA	TGAGGAAGGA	TTGGAGAGCC
721	ACTTGAGGTC	ATATGTTAAG	TACGCGTATA	AGGCTGAGCC	ATATGTTGCC	TCTGAATACA	AGACAGTGCA	TGAAGAAGCT
801	ACCAAATCCA	TGACCACGAT	TCTCAAGCCT	TCTGCCGATT	TCCTCACCAG	CAACAACTA	CTGAGGTACT	CATGGTTTTT
881	CTTTGATGTA	CTGATCAAA	CTATGGCTCA	GCATTGATA	GAGAACTCCA	AAGTTAAGTT	GCTGCGAAAC	CAGAGATTTT
961	CTGCATCCTA	TCATCATGCA	GCGGAAACCG	TTGTAATAT	GCTGATGCCA	CACATCACTC	AGAAGTTTGG	AGATAATCCA
1041	GAGGCATCTA	AGAACGCGAA	TCATAGCCTT	GCTGTCTTCA	TCAAGAGATG	TTTCACCTTC	ATGGACAGGG	GCTTTGTCTT
1121	CAAGCAGATC	AACAACCTACA	TTAGCTGTTT	GCTCCTGGA	GACCCAAAGA	CCCTCTTGA	ATACAAGTTT	GAATTTCTCC
1201	GTGTAGTGTG	CAACCATGAA	CATTATATTC	CGTTGAACTT	ACCAATGCCA	TTTGGAAAAG	GCAGGATTCA	AAGATACCAA
1281	GACCTCCAGC	TTGACTACTC	ATTAACAGAT	GAGTTCTGCA	GAAACCACTT	CTTGGTGGGA	CTGTTACTGA	GGGAGGTGGG
1361	GACAGCCCTC	CAGGAGTTCC	GGGAGGTCCG	TCTGATCGCC	ATCAGTGTGC	TCAAGAACCT	GCTGATAAAG	CATTCTTTTG
1441	ATGACAGATA	TGCTTCAAGG	AGCCATCAGG	CAAGGATAGC	CACCTCTAC	CTGCTCTGT	TTGGTCTGCT	GATTGAAAAC
1521	GTCCAGCGGA	TCAATGTGAG	GGATGTGTCA	CCCTTCCCTG	TGAACGCGGG	CATGACCGTG	AAGGATGAAT	CCCTGGCTCT
1601	ACCAGCTGTG	AATCCGCTGG	TGACGCGCGA	GAAGGGAAGC	ACCCTGGACA	ACAGCCTGCA	CAAGGACCTG	CTGGGCGCTC
1681	TCTCCGGCAT	TGCTTCTCCA	TATACAACCT	CAACTCCAAA	CATCAACAGT	GTGAGAAATG	CTGATTCTGAG	AGGATCTCTC
1761	ATAAGCACAG	ATTCCGGTAA	CAGCCTTCCA	GAAAGGAATA	GTGAGAAGAG	CAATTCCCTG	GATAAGCACC	AACAAAGTAG
1841	CACATTGGGA	AATTCCGTGG	TTCCGTGTGA	TAAACTTGAC	CAGTCTGAGA	TTAAGAGCCT	ACTGATGTGT	TTCTCTTACA
1921	TCTTAAAGAG	CATGTCTGAT	GATGCTTTGT	TTACATATTG	GAACAAGGCT	TCAACATCTG	AACCTATGGA	TTTTTTTACA
2001	ATATCTGAAG	TCTGCCTGCA	CCAGTTCCAG	TACATGGGGA	AGCGATACAT	AGCCAGGAAC	CAGGAGGGGT	TGGGACCCAT
2081	AGTTTCATGAT	CGAAAGTCTC	AGACATTGCC	TGTTTCCCGT	AACAGAACAG	GAATGATGCA	TGCCAGATTG	CAGCAGCTGG
2161	GCAGCCTGGA	TAACCTCTCT	ACTTTTAAAC	ACAGCTATGG	CCACTCGGAC	CAGATGTTTC	TGCACCAGTC	ATTACTTGAA
2241	GCCAACATTG	CTACTGAGGT	TGCTCTGACA	GCTCTGGACA	CGCTTTCTCT	ATTTACATTG	GCGTTTAAAG	ACCAGCTCCT
2321	GGCCGACCAT	GGACATAATC	CTCTCATGAA	AAAAGTTTTT	GATGTCTACC	TGTGTTTTCT	TCAAAAACAT	CAGTCTGAAA
2401	CGGCTTTAAA	AAATGTCTTC	ACTGCCTTAA	GGTCCCTAAT	TTATAAGTTT	CCCTCAACAT	TCTATGAAGG	GAGAGCGGAC
2481	ATGTGTGCGG	CTCTGTGTTA	CGAGATTCTC	AAGTGCTGTA	ACTCCAAGCT	GAGCTCCATC	AGGACGGAGG	CCTCCAGCTC
2561	GCTCTACTTC	CTGATGAGGA	ACAACCTTGA	TTACACTGGA	AAGAAGTCCT	TTGTCCGGAC	ACATTTGCAA	GTCATCATAT
2641	CTGTTCAGCA	GCTGATAGCA	GACGTGTGTG	GCATTGGGGA	AACCAGATTG	CAGCAGTCCC	TGTCCATCAT	CAACAACGTG
2721	GCCAACAGTG	ACCGGCTTAT	TAAGCACACC	AGCTTCTCCT	CTGATGTGAA	GGACTTAACC	AAAAGGATAC	GCACGGTGCT
2801	AATGGCCACC	GCCCAGATGA	AGGAGCATGA	GAACGACCCA	GAGATGCTGG	TGGACCTCCA	GTACAGCCTG	GCCAAATCCT
2881	ATGCCAGCAC	GCCCAGCTTC	AGGAAGACGT	GGCTCGACAG	CATGGCCAGG	ATCCATGTCA	AAAATGGCGA	TCTCTCAGAG
2961	GCAGCAATGT	GCTATGTCCA	CGTAACAGCC	CTAGTGGCAG	AATATCTCAC	ACGGAAGGCG	GTGTTTAGAC	AAGGATGCAC
3041	CGCTTTCAGG	GTCATTACCC	CAAACATCGA	CGAGGAGGCC	TCCATGATGG	AAGACGTGGG	GATGCAGGAT	GTCCATTTCA
3121	ACGAGGATGT	GCTGATGGAG	CTCCTTGAGC	AGTGGCGAGA	TGGACTCTGG	AAAGCCGAGC	GCTACGAGCT	CATCGCCGAC
3201	ATCTACAAAC	TTATCATCCC	CATTTATGAG	AAGCGGAGGG	ATTTCTTTGA	AGATGAAGAT	GGAAAGGAGT	ATATTTACAA
3281	GGAACCCAAA	CTCACCCGCA	TGTCGGAAAT	TTCTCAGAGA	CTCCTTAAAC	TGTACTCGGA	TAAATTTGGT	TCTGAAAATG
3361	TCAAAATGAT	ACAGGATTCT	GGCAAGGTCA	ACCCTAAGGA	TCTGGATTCT	AAGTATGCAT	ACATCCAGGT	GACTCACGTC
3441	ATCCCTTCT	TTGACGAAAA	AGAGTTGCAA	GAAAGGAAAA	CAGAGTTTGA	GAGATCCAC	AACATCCGCC	GCTTCATGTT
3521	TGAGATGCCA	TTTACGCAGA	CCGGGAAGAG	GCAGGGCGGG	GTGGAAGAGC	AGTGCAAACG	GCGCACCATC	CTGACAGCCA
3601	TACACTGCTT	CCCTTATGTG	AAGAAGCGCA	TCCCTGTCTAT	GTACCAGCAC	CACACTGACC	TGAACCCCAT	CGAGGTGGCC
3681	ATTGACGAGA	TGAGTAAGAA	GGTGGCGGAG	CTCCGGCAGC	TGTGCTCCTC	GGCCGAGGTG	GACATGATCA	AACTGCAGCT
3761	CAAACCTCCAG	GGCAGCGTGA	GTGTTTCAGG	CAATGCTGGC	CCACTAGCAT	ATGCGCGAGC	TTTCTTAGAT	GATACAAACA
3841	CAAAGCGATA	TCCTGACAAT	AAAGTGAAGC	TGCTTAAGGA	AGTTTTTCAGG	CAATTTGTGG	AAGCTTGCGG	TCAAGCCTTA
3921	GCGGTAAACG	AACGTCTGAT	TAAAGAAGAC	CAGCTCGAGT	ATCAGGAAGA	AATGAAAGCC	AACTACAGGG	AAATGGCGAA
4001	GGAGCTTTCT	GAAATCATGC	ATGAGCAGAT	CTGCCCTCTG	GAGGAGAAGA	CGAGCGTCTT	ACCGAATTCC	CTTCACATCT
4081	TCAACGCCAT	CAGTGGGACT	CCAACAAGCA	CAATGGTTCA	CGGGATGACC	AGCTCGTCTT	CGGTCTGTGT	ATTACATCTC
4161	ATGGCCCGTG	TGTGGGGACT	TGCTTTGTCA	TTTGCAAACT	CAGGATGCTT	TCCAAAGCCA	ATCACTGGGG	AGACCGAGCA
4241	CAGGGAGGAC	CAAGGGGAAG	GGGAGAGAAA	GGAAATAAAG	AACAACGTTA	TTTCTTAACA	GACTTTCTAT	AGGAGTTGTA
4321	AGAAGGTGCA	CATATTTTTT	TAAATCTCAC	TGGCAATATT	CAAAGTTTTT	ATTGTGTCTT	AACAAAGGTG	TGGTAGACAC
4401	TCTTGAGCTG	GACTTAGATT	TTATTCTTCC	TGTCAGAGTA	GTGTTAGAAT	AGATGGCCTA	CAGAAAAAAA	AGGTTCTGGG
4481	ATCTACATGG	CAGGGAGGGC	TGCACTGACA	TTGATGCCCTG	GGGGACCTTT	TGCCCTGACT	CGTGCCGGAA	ATCTGATCGT
4561	AATCAGGGTA	CAGAACTTAC	TAGTTTTGTC	TAGGAGTATG	TTGTATGACT	AGGATTTGTG	CTATTATCTC	ATTCAACAAC
4641	ATAGAGCAAG	AATAGTGAGC	TAACTGAGCT	AGACACTCAA	TTAATCCGCT	ACTGGCTTCA	AGTCAGAACT	TTGTCAATTA
4721	TCATCGACTC	CGGGACGGTC	ATATATGTAT	TACATTTCTA	CATTTTTAAT	ACTCACATGG	GCTTATGCAT	TAAGTTTAAT
4801	TGTGATAAAT	TTGTGCTGGT	CCAGTATATG	CAATACACTT	TAATGGTTTA	TTCTTGTCTAT	AAAAATGTGC	AATATGGAGA
4881	TGTATACAAG	TCTTTACT						

Fig. 10B



	10	20	30	40	50	60	70	80
1	MEGHVMIAFL	PTILNQLFRV	LTRATQEEVA	VNVTRVIIHV	VAQCHEEGLE	SHLRSYVKYA	YKAEFVASE	YKTVHEELTK 80
81	SMTIILKPSA	DFLTSTNKLRL	YSWFFFDVLI	KSMAQHLIEN	SKVKLLRNQR	FPASYHHAAE	TVVNMMLPHI	TQKFGDNPEA 160
161	SKNANHSLAV	FIKRCFTFMD	RGFVFKQINN	YISCFAPGDP	KTLFEYKFEF	LRVVCNHEHY	IPLNLPMPFG	KGRIQRYQDL 240
241	OLDYSLTDEF	CRNHFLVGLL	LREVGTAQNE	FREVRLIAIS	VLKXNLLIKHS	FDDRYASRSH	QARIATLYLP	LFGLLIENVQ 320
321	RINVRDVSPF	PVNAGMTVKD	ESLALPAVNP	LVTPOKGSTL	DNSLHKDLLG	AISGIASPYT	TSTPNINSVR	NADSRGSLIS 400
401	TDSGNSLPER	NSEKSNSLDK	HQSSSTLGNS	VVRCDKLDQS	EIKSLMCFL	YILKSMSDDA	LFTYWNKAST	SELMDDFTIS 480
481	EVCLHQFOYM	GKRYIARNQE	GLGPIVHDRK	SOTLFSVRNR	TGMMHARLQQ	LGSLDNSLTF	NHSYGHSDAD	VLHQSLLEAN 560
561	IATEVCLTAL	DTLSLFTLAF	KNQLLADHGH	NPLMKKVFDV	YLCFLQKHQS	ETALKNVFTA	LRSLIYKFPS	TFYEGRADMC 640
641	AALCYEILKC	CNSKLSIRT	EASQLLYFLM	RNNFDYTGKK	SFVRTHLQVI	ISVSQLIADV	VGIGETRFOQ	SLSIINNCAN 720
721	SDRLIKHTSF	SSDVKDLTRK	IRTVLMATAQ	MKEHENDPEM	LVDLQYSLAK	SYASTPELRK	TWLDSMARIH	VKNGDLSEAA 800
801	MCYVHVITALV	AEYLTRKGVP	RQGCTAFRVI	TPNIDEEASM	MEDVGMQDVH	FNEDVLMELL	EQCADGLWKA	ERYELIADIY 880
881	KLIPIIYEKR	RDFFEDEDGK	EYIYKEPKLT	PLSEISORLL	KLYSDKPGSE	NVKMIQDSGK	VNPKDLDSKY	AYIQVTHVIP 960
961	FFDEKELQER	KTEFERSHNI	RRPMFEMPFT	QTGKRQGGVE	EQCKRRTILT	AIHCFPYVKK	RIPVMYQHHT	DLNPIEVAID 1040
1041	EMSKKVAELR	QLCSSAEVDM	IKLQKLQGS	VSVQVNAGPL	AYARAFLDDT	NTKRYPDNKV	KLLKEVFRRQF	VEACGQALAV 1120
1121	NERLIKEDQL	EYQEEMKANY	REMAKELSEI	MHEQICPLEE	KTSVLPNSLH	IFNAISGTP	STMVHGMTSS	SSVV 1194
	10	20	30	40	50	60	70	80

Figure 1 displays 12 histograms showing the distribution of the number of non-zero elements in the vector  $x$  for different values of  $n$  (10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120). The x-axis represents the number of non-zero elements ( $x$ ), and the y-axis represents the count. As  $n$  increases, the distribution becomes more concentrated around zero, indicating that the vector  $x$  is becoming sparser.

FIG. 10B (cont.)

	10	20	30	40	50	60	70	80
1	AATTGTAATA	CGACTCACTA	TAGGGCGAAT	TGGGTACCGG	GCCCCCCTC	GAGGTCGACG	GTATCGATAA	GCTTGATATC
2	GAATTCGGCA	CGAGTTTTAC	ACCATCACCA	AAACCCAGAA	TTTTATGATG	AGATTAAAT	AGAGTTGCC	ACTCAGCTGC
3	ATGAAAAGCA	CCACCTGTTG	CTCACATTCT	TCCATGTCAG	CTGTGACAAC	TCAAGTAAAG	GAAGCACGAA	GAAGAGGGAT
4	GTCGTTGAAA	CCCAAGTTGG	CTACTCCTGG	CTTCCCCTCC	TGAAAGACGG	AAGGGTGGTG	ACAAGCGAGC	AGCACATCCC
5	GGTCTCGGCG	AACCTTCCTT	CGGGCTATCT	TGGCTACCAA	GAGCTTGGGA	TGGGCAGGCA	TTATGGTCCG	GAAATTAAAT
6	GGGTAGATGG	AGGCAAGCCA	CTGCTGAAAA	TTTCCACTCA	TCTGGTTTCT	ACAGGGATAC	TCAGGATCAG	CATTTACATA
7	ATTTTTTCCA	GTACTGTCAG	AAAACCGAAT	CTGGAGCCCA	AGCCTTAGGA	AACGAACCTG	TAAAGTACCT	TAAGAGTCTG
8	CATGCGATGG	AAGGCCACGT	GATGATCGCC	TTCTTGCCCA	CTATCCTAAA	CCAGCTGTTT	CGAGTCTCTA	CCAGAGCCAC
9	ACAGGAAGAA	GTGCGGGTTA	ACGTGACTCG	GGTCATTATT	CATGTGGTTG	CCCAGTGCCA	TGAGGAAGGA	TTGGAGAGCC
10	ACTTGAGGTC	ATATGTTAAG	TACGCGTATA	AGGCTGAGCC	ATATGTTGCC	TCTGAATACA	AGACAGTGCA	TGAAGAACTG
11	ACCAAATCCA	TGACCACGAT	TCTCAAGCCT	TCTGCCGATT	TCCTCACCAG	CAACAACTA	CTGAGGTACT	CATGGTTTTT
12	CTTTGATGTA	CTGATCAAAT	CTATGGCTCA	GCAATTTGATA	GAGAACTCCA	AAGTTAAGTT	GCTGCGAAAC	CAGAGATTTT
13	CTGCATCCTA	TCATCATGCA	GCGGAAACCG	TTGTAAATAT	GCTGATGCCA	CACATCAGTC	AGAAGTTTGG	AGATAATCCA
14	GAGGCATCTA	AGAACGCGAA	TCATAGCCTT	GCTGTCTTCA	TCAAGAGATG	TTTCACTTTC	ATGGACAGGG	GCTTTGTCTT
15	CAAGCAGATC	AACAACATCA	TTAGCTGTTT	TGCTCCTGGA	GACCCAAAGA	CCCTCTTTEA	ATACAAGTTT	GAATTTCTCC
16	GTGTAGTGTG	CAACCATGAA	CATTATATTC	CGTTGAACTT	ACCAATGCCA	TTTGGAAAAG	GCAGGATTCA	AAGATACCAA
17	GACCTCCAGC	TTGACTACTC	ATTAACAGCT	GAGTCTGCA	GAAACCACTT	CTTGGTGGGA	CTGTTACTGA	GGGAGGTGGG
18	GACAGCCCTC	CAGGAGTTCC	GGGAGGTCCG	TCTGATCGCC	ATCAGTGTGC	TCAAGAACCT	GCTGATAAAG	CATTCTTTTG
19	ATGACAGATA	TGCTTCAAGG	AGCCATCAGG	CAAGGATAGC	CACCTCTAC	CTGCTCTGCT	TTGGTCTGCT	GATTGAAAAC
20	GTCCAGCGGA	TCAATGTGAG	GGATGTGTCA	CCCTTCCCTG	TGAACGCGGG	CATGACCGTG	AAGGATGAAT	CCCTGGCTCT
21	ACCAGCTGTG	AATCCGCTGG	TGACGCCGCA	GAAGGGAAGC	ACCCTGGACA	ACAGCCTGCA	CAAGGACCTG	CTGGGCGCCA
22	TCTCCGGCAT	TGCTTCTCCA	TATACAACCT	CAACTCCAAA	CATCAACAGT	GTGAGAAATG	CTGATTCCAG	AGGATCTCTC
23	ATAAGCACAG	ATTCCGGTAA	CAGCCTTCCA	GAAAGGAATA	GTGAGAAGAG	CAATTCCCTG	GATAAGCACC	AACAAAGTAG
24	CACATTGGGA	AATTCCTGGG	TTCCGCTGTA	TAAACTTGAC	CAGTCTGAGA	TTAAGAGCCT	ACTGATGTGT	TTCTCTTACA
25	TCTTAAAGAG	CATGCTGTAT	GATGCTTTGT	TTACATATTG	GAACAAGGCT	TCAACATCTG	AACCTATGGA	TTTTTTTACA
26	ATATCTGAAG	TCTGCCCTGA	CCAGTTCCAG	TACATGGGGA	AGCGATACAT	AGCCAGGAAC	CAGGAGGGGT	TGGGACCCAT
27	ATTCATGAT	CGAAAGTCTC	AGACATTGCC	TGTTTCCCGT	AACAGAACAG	GAATGATGCA	TGCCAGATTG	CAGCAGCTGG
28	AGAGCCTGGA	TAACTCTCTC	ACTTTTAAAC	ACAGCTATGG	CCACTCGGAC	GCAGATGTTT	TGCACCAGTC	ATTACTTGAA
29	GCCAACATTG	CTACTGAGGT	TTGCCCTGACA	GCTCTGGAGA	CGCTTCTCTT	ATTTACATTG	GCGTTTAAAG	ACCAGCTCCT
30	GGCCGACCAT	GGACATAATC	CTCTCATGAA	AAAAGTTTTT	GATGTCTACC	TGTGTTTTCT	TCAAAAACAT	CAGTCTGAAA
31	CGGCTTTAAA	AAATGTCTTC	ACTGCCTTAA	GGTCCTTAAT	TTATAAGTTT	CCCTCAACAT	TCTATGAAGG	GAGAGCGGAC
32	ATGTGTGCGG	CTCTGTGTTA	CGAGATTCTC	AAGTGCTGTA	ACTCCAAGCT	GAGCTCCATC	AGGACGGAGG	CCTCCAGCT
33	GCTCTACTTC	CTGATGAGGA	ACAACCTTGA	TTACACTGGA	AAGAAGTCTT	TTGTCCGGAC	ACATTTGCAA	GTATCATAT
34	CTGTACAGCA	ACGTGTTGTT	GACGTTGTTG	GCATTGGGGA	AACCAGATTG	CAGCAGTCCC	TGTCCATCAT	CAACAACCTG
35	GCCAACAGTG	ACCGGCTTAT	TAAGCACACC	AGCTTCTCCT	CTGATGTGAA	GGACTTAAAC	AAAAGGATAC	GCACGGTGCT
36	ATGGCCACC	GCCCAGATGA	AGGAGCATGA	GAACGACCCA	GAGATGCTGG	TGGACCTCCA	GTACAGCCTG	GCCAAATCCT
37	ATGCCAGCAC	GCCCAGAGTC	AGGAAGACGT	GGCTCGACAG	CATGGCCAGG	ATCCATGTCA	AAAATGGCGA	TCTCTCAGAG
38	GCAGCAATGT	GCTATGTCCA	CGTAACAGCC	CTAGTGGCAG	AATATCTCAC	ACGGAAAGGC	GTGTTTAGAC	AAGGATGCAC
39	CGCCTTCAGG	GTCATTACCC	CAAACATCGA	CGAGGAGGCC	TCCATGATGG	AAGACGTGGG	GATGCAGGAT	GTCCATTTC
40	ACGAGGATGT	GCTGATGGAG	CTCCTTGAGC	AGTGCGCAGA	TGGACTCTGG	AAAGCCGAGC	GCTACGAGCT	CATGCGGAC
41	ATCTACAAAC	TTATCATCCC	CATTTATGAG	AAGCGGAGGG	ATTCTTTTGA	AGATGAAGAT	GGAAAGGAGT	ATATTTACAA
42	GGAAACCCAA	CTCACACCGC	TGTCGGAAAT	TTCTCAGAGA	CTCCTTAAAC	TGTACTCGGA	TAAATTTGGT	TCTGAAAATG
43	TCAAAATGAT	ATCAGATTCT	GGCAAGGTCA	ACCCTAAGGA	TCTGGATTCT	AAGTATGCAT	ACATCCAGGT	GACTCACGTC
44	ATCCCTTTCT	TTGACGAAAA	AGAGTTGCAA	GAAAGGAAAA	CAGAGTTTGA	GAGATCCCAC	AACATCCGCC	GCTTCATGTT
45	TGAGATGCCA	TTTACGCGAA	CCGGGAAGAG	GCAGGGCGGG	GTGGAAGAGC	AGTGCAAACG	GCGCACCATC	CTGACAGCCA
46	TACACTGCTT	CCCTTATGTG	AAGAAGCGCA	TCCCTGTCTAT	GTACCAGCAC	CACACTGACC	TGAACCCCAT	CGAGGTGGCC
47	ATTGACGAGA	TGAGTAAGAA	GGTGGCGGAG	CTCCGGCAGC	TGTGCTCCTC	GGCCGAGGTG	GACATGATCA	AACTGCAGCT
48	CAAACCTCCAG	GGCAGCGTGA	GTGTTCAAGT	CAATGCTGGC	CCACTAGCAT	ATGCGCGAGC	TTTCTTAGAT	GATACAAACA
49	CAAAGCGATA	TCCTGACAAT	AAAGTGAAGC	TGCTTAAAGGA	AGTTTTTCAGG	CAATTTGTGG	AAGCTTGGCG	TCAAGCCTTA
50	GCGGTAAACG	AACGTCTGAT	TAAAGAAGAC	CAGCTCGAGT	ATCAGGAAGA	AATGAAAGCC	AACTACAGGG	AAATGGCGAA
51	GGAGCTTTCT	GAAATCATGC	ATGAGCAGAT	CTGCCCCCTG	GAGGAGAAGA	CGAGCGTCTT	ACCGAATTCC	CTTCACATCT
52	TCAACGCCAT	CAGTGGGACT	CCAACAAGCA	CAATGGTTCA	CGGGATGACC	AGCTCGTCTT	CGGTGCTGTG	ATTACATCTC
53	ATGGCCCGTG	TGTGGGGACT	TGCTTTGTCA	TTTGCAAACT	CAGGATGCTT	TCCAAAGCCA	ATCACTGGGG	AGACCGAGCA
54	CAGGGAGGAC	CAAGGGGAAG	GCGAGAGAAA	GGAAATAAAG	AACAACGTTA	TTTCTTAAAC	GACTTCTTAT	AGGAGTTGTA
55	AGAAGGTGCA	CATATTTTTT	TAAATCTCAC	TGGCAATATT	CAAAGTTTTT	ATTGTGCTCT	AACAAAGGTG	TGGTAGACAC
56	TCTTGAGCTG	GACTTAGATT	TTATTCTTCC	TTGCAGAGTA	GTGTTAGAAT	AGATGGCCTA	CAGAAAAAAA	AGGTTCTGGG
57	ATCTACATGG	CAGGGAGGGC	TGCACTGACA	TTGATGCCTG	GGGGACCTTT	TGCCCTGACT	CGTGCCGGAA	ATCTGATCGT
58	AATCAGGGTA	CAGAACTTAC	TAGTTTTGTC	TAGGAGTATG	TTGTATGACT	AGGATTTGTG	CTATTATCTC	ATTCAACAAC
59	ATAGAGCAAG	AATAGTGAGC	TAACTGAGCT	AGACACTCAA	TTAATCCGCT	ACTGGCTTCA	AGTCAGAACT	TTGTCAATTA
60	TCATCGACTC	CGGGACGGTC	ATATATGTAT	TACATTTCTA	CATTTTTAAT	ACTCACATGG	GCTTATGCTT	TAAAGTTAAT
61	TGTGATAAAT	TTGTGCTGGT	CCAGTATATG	CAATACACTT	TAATGGTTTA	TTCTTGTCAT	AAAAATGTGC	AATATGGAGA
62	TGTATACAAG	TCTTTACT						

Fig. 10C

	10	20	30	40	50	60	70	80
MEGHVMI AFL	PTILNQLFRV	LTRATOE EVA	VNVTRV IIV	VAQCHEEGLE	SHLSYV KYA	YKAE PYVASE	YKTVHEELTK	80
SMTTILKPSA	DFLT SNKLLR	YSWFFFDVLI	KSMAOHLIEN	SKVKLLRNOR	FPASYHHAEE	TVVNMLMPHI	TOKFGDNPEA	160
SKNANHSLAV	FIKRCPTFMD	RGFVFKQINN	YISC FAPGDP	KTLFEYKFEP	LRVVCNHEHY	IPLNLPMPFG	KGRIORYQDL	240
QLDYSLTDEF	CRNHFLVGLL	LREVG TALQE	FREVRLIAIS	VJKNLLIKHS	FDDRYASRSH	QARIATLYLF	LPGLLIENVO	320
RINVRDVSPF	PVNAGMTVKD	ESLALPAVNP	LVTPOKGSTL	DNSLHKOLLG	AISGLASPYT	TSTPNINSVR	NADSRGSLIS	400
TDSGNSLPER	NSEKNSLSDK	HQOSSTLGNS	VVRCDKLDQS	EIKSLMCFI	YILKSMSDDA	LFTYWNKAST	SELMOFFTIS	480
EVCLHQFOYM	GKRYIARNQE	GLGPIVHDRK	SQTLPVSRNR	TGMTHARLOQ	LGSLDNSLTF	NHSYGHSDAD	VLHQSLLEAN	560
IATEVCLTAL	DTLSLFTLAF	KNQLLADHGH	NPLMKRVFDV	YLCFLOKHQS	ETALKNVFTA	LRSLTYKFPS	TFYEGRADMC	640
AALCYEILKC	CNSKLSIRT	EASOLLYFLM	RNNFDYTGK	SFVRTHLOVI	ISVSQLIADV	VGIGETRPQQ	SLSIINN CAN	720
SDRLIKHTSF	SSDVKDLTKR	IRTVLMATAQ	MKEHENDPEM	LVDLOYS LAK	SYASTPELRK	TWLD SMARH	VKNGDLSEAA	800
MCYVHV TALV	AEYLTRKGVF	ROGCTAFRVI	TPNIDEEASM	MEDVGMQDVH	FNEDVLMELL	EOCADGLMKA	ERYELIADIY	880
KLIPIYEKR	RDFFEDEDGK	EYTYKEPKLT	PLSEISORLL	KLYSDKFGSE	NVKMIQDSGK	VNPKDLD SKY	AYIQVTHVIP	960
FFDEKELQER	KTEFERSHNI	RRFMFEMPFT	QTGKRQGGVE	EQCKERTILT	AIHCFFPYVK	RIPVMYQHHT	DIANPIEVAID	1040
EMSKKVAELR	QLC SSAEVDN	IKLOKLQGS	VSVQVNAGPL	AYARAF LDDT	NTKRYFDNKV	KLLKEVFRQF	VEACQALAV	1120
NERLIKEDQL	EYQEE MKANY	REMAKELSEI	MHEQICPLEE	KTSVLPNSLH	IFNAISGTPT	STMVHGMTSS	SSVV	1194

FIG. 10C (cont.)

	10	20	30	40	50	60	70	80
1	AATTGTAATA	CGACTCACTA	TAGGGCGAAT	TGGGTACCGG	GCCCCCTCTC	GAGGTGCGACG	GTATCGATAA	GCTTGATATC
81	GAATTCGGCA	CGAGTTTAC	ACCATCACCA	AAACCCAGAA	TTTTATGATG	AGATTAAAT	AGAGTTGCC	ACTCAGCTGC
61	ATGAAAAGCA	CCACCTGTTG	CTCACATTCT	TCCATGTCAG	CTGTGACAAC	TCAAGTAAAG	GAAGCACGAA	GAAGAGGGAT
41	GTCGTTGAAA	CCCAAGTTGG	CTACTCCTGG	CTTCCCCCTCC	TGAAAGACGG	AAGGGTGGTG	ACAAGCGAGC	AGCACATCCC
21	GGTCTCGGCG	AACCTTCCTT	CGGGCTATCT	TGGCTACCAA	GAGCTTGGGA	TGGGCAGGCA	TTATGGTCCG	GAAATTAAAT
01	GGGTAGATGG	AGGCAAGCCA	CTGCTGAAAA	TTTCCACTCA	TCTGGTTTCT	ACAGGGATAC	TCAGGATCAG	CATTTACATA
81	ATTTTTTCCA	GTACTGTCAG	AAAACCGAAT	CTGGAGCCCA	AGCCTTAGGA	AACGAACTTG	TAAAGTACCT	TAAGAGTCTG
61	CATGCGATGG	AAGGCCACGT	GATGATCGCC	TTCTTGCCCA	CTATCCTAAA	CCAGCTGTTT	CGAGTCCCTCA	CCAGAGCCAC
41	ACAGGAAGAA	GTGCGGGTTA	ACGTGACTCG	GGTCATTATT	CATGTGGTTG	CCCAGTGCCA	TGAGGAAGGA	TTGGAGAGCC
21	ACTTGAGGTC	ATATGTTAAG	TACGCGTATA	AGGCTGAGCC	ATATGTTGCC	TCTGAATACA	AGACAGTGCA	TGAAGAACTG
01	ACCAAATCCA	TGACCACGAT	TCTCAAGCCT	TCTGCGGATT	TCCTCACCAG	CAACAACTA	CTGAGGTACT	CATGGTTTTT
881	CTTTGATGTA	CTGATCAAAT	CTATGGCTCA	GCATTTGATA	GAGAACTCCA	AAGTTAAGTT	GCTGCGAAAC	CAGAGATTTT
962	CTGCATCCTA	TCATCATGCA	GCGGAAACCG	TTGTAAATAT	GCTGATGCCA	CACATCACTC	AGAAGTTTGG	AGATAATCCA
041	GAGGCATCTA	AGAACGCGAA	TTATAGCCTT	GCTGTCTTCA	TCAAGAGATG	TTTCACTTTC	ATGGACAGGG	GCTTTGTCTT
121	CAAGCAGATC	AACAACCTACA	CATTATGTTT	TGCTCCTGGA	GACCCAAAGA	CCCTCTTTGA	ATACAAGTTT	GAATTTCTCC
201	GTGTAGTGTG	CAACCATGAA	CATTATATTC	CGTTGAACTT	ACCAATGCCA	TTTGGAAGAG	GCAGGATTCA	AAGATACCAA
1281	GACCTCCAGC	TTGACTACTC	ATTAACAGAT	GAGTTCTGCA	GAAACCACTT	CTTGGTGGGA	CTGTTACTGA	GGGAGGTGGG
1361	GACAGCCCTC	CAGGAGTTCC	GGGAGGTCCG	TGTGATCGCC	ATCAGTGTGC	TCAAGAACCT	GCTGATAAAG	CATTCCTTTG
1441	ATGACAGATA	TGCTTCAAGG	AGCCATCAGG	CAAGGATAGC	CACCTCTAC	CTGCCCTCTG	TTGGTCTGCT	GATTGAAAAC
1521	GTCCAGCGGA	TCAATGTGAG	GGATGTGTCA	CCCTTCCCTG	TGAACGCGGG	CATGACCTG	AAGGATGAAT	CCCTGGCTCT
1601	ACCAGCTGTG	AATCCGCTGG	TGACGCGGCA	GAAGGGAAGC	ACCCTGGACA	ACAGCCTGCA	CAAGGACCTG	CTGGGCGCCA
1681	TCTCCGGCAT	TGCTTCTCCA	TATACAACCT	CAACTCCAAA	CATCAACAGT	GTGAGAAATG	CTGATTGAG	AGGATCTCTC
1761	ATAAGCACAG	ATTCCGGTAA	CAGCCTTCCA	GAAAGGAATA	GTGAGAAGAG	CAATTCCTTG	GATAAGCACC	AACAAAGTAG
1841	CACATTGGGA	AATTCOGTGG	TTGCTGTGTA	TAAACTTGAC	CAGTCTGAGA	TTAAGAACCT	ACTGATGTGT	TTCTCTTACA
1921	TCTTAAAGAG	TCTGCTTGCA	CCAGTTCAG	TACATGGGGA	GAACAAGGCT	TCAACATCTG	AACCTATGGA	TTTTTTTACA
2001	ATATCTGAAG	CGAAAGTCTC	AGACATTGCC	TGTTTTCCGT	AGCGATACAT	AGCCAGGAAC	CAGGAGGGGT	TGGGACCCAT
2081	AGTTCATGAT	TAACCTCTCT	ACTTTTAAAC	ACAGCTATGG	AACAGAACAG	GAATGATGCA	TGCCAGATTG	CAGCAGCTGG
2161	GCAGCCTGGA	CTACTGAGGT	TTGCCCTGACA	ACCTCTGGACA	CCACTCGGAC	GCAGATGTTT	TGCACCAGTC	ATTACTTGAA
2241	GCCAACATTG	GGACATAATC	CTCTCATGAA	AAAAGTTTTT	CGCTTTCTCT	ATTTACATTT	GCGTTTAAGA	ACCAGCTCCT
2321	GGCCGACCAT	AAATGCTCTC	ACTGCCCTTAA	GGTCCCTAAT	GATGTCTACC	TGTGTTTTCT	TCAAAAACAT	CAGTCTGAAA
2401	GGGCTTTAAA	CTCTGTGTTA	CGAGATTCTC	AAGTGTCTGT	TTATAAGTTT	CCCTCAACAT	TCTATGAAGG	GAGAGCGGAC
2481	ATGTGTGCGG	CTGATGAGGA	ACAACCTTGA	TTACACTGGA	ACTCCAAGCT	GAGCTCCATC	AGGACGGAGG	CTCCCGAGCT
2561	GCTCTACTTC	CGTGATAGCA	GACGTTGTTG	GCATTGGGGA	AAGAAGTCTC	TTGTCCGGAC	ACATTTGCAA	GTATCATAT
2641	GTGTAGGCA	ACCGGCTTAT	TAAGCACACC	AGCTTCTCCT	AACCAGATTC	CAGCAGTCCC	TGTCCATCAT	CAACAACCTG
2721	GCCAACAGTG	GCCCAGATGA	AGGAGCATGA	GAACGACCCA	CTGATGTGAA	GGACTTAACC	AAAAGGATAC	GCACGGTGCT
2801	AATGGCCACC	GCCCAGAGTC	AGGAAGACGT	GGCTCGACAG	GAGATGCTGG	TGGACCTCCA	GTACAGCCTG	GCCAAATCCT
2881	ATGCCAGCAC	GCTATGTCCA	CGTAACAGCC	CGAGGAGGCC	CATGGCCAGG	ATCCATGTCA	AAAATGGOGA	TCTCTCAGAG
2961	GCAGCAATGT	GTCATTACCC	CAAACATCGA	AGTGCGCAGA	AATATCTCAC	ACGGAAGGCG	GTGTTTAGAC	AAGGATGCAC
3041	CGCCTTCAGG	GCTGATGGAG	CTCCTTGAGC	AAGCGGAGGG	TCCATGATGG	AAGACGTGGG	GATGCAGGAT	GTCCATTTCA
3121	ACGAGGATGT	TTATCATCCC	CATTTATGAG	TGTCGGAAT	TGGACTCTGG	AAAGCCGAGC	GCTACGAGCT	CATCGCCGAC
3201	ATCTACAAAC	CTCACACCGC	TGTCGGAAT	TTCTCAGAGA	ATTTCTTTGA	AGATGAAGAT	GGAAAGGAGT	ATATTTACAA
3281	GGAACCCAAA	ACAGGATTCT	GGCAAGGTCA	ACCCTAAGGA	CTCCTTAAAC	TGTACTCGGA	TAAATTTGGT	TCTGAAAATG
3361	TCAAAATGAT	TTGACGAAAA	AGAGTTGCAA	GAAAGGAAAA	TCTGGATTCT	AAGTATGCAT	ACATCCAGGT	GACTCACGTC
3441	ATCCCCCTCT	TTTACGCGA	CCGGAAGAG	GCAGGGCGGG	CAGAGTTTGA	GAGATCCCAC	AACATCCGCG	GCTTCATGTT
3521	TGAGATGCCA	CCCTTATGTG	AAGAAGCGCA	TCCCTGTCTAT	GTGGAAGAGC	AGTGCAAACG	GCGCACCATC	CTGACAGCCA
3601	TACACTGCTT	GGCAGCGTGA	GTGTTCAAGT	CAATGCTGGC	GTACCAGCAC	CACACTGACC	TGAACCCCAT	CGAGGTGGCC
3681	ATTGACGAGA	TCCTGACAAT	AAAGTGAAGC	TGCTTAAAGGA	TGTGCTCCTC	GGCCGAGGTG	GACATGATCA	AACTGCAGCT
3761	CAAACCTCCAG	AACGTCTGAT	TAAAGAAGAC	CAGCTCGAGT	CCACTAGCAT	ATGCGCGAGC	TTTCTTAGAT	GATACAAACA
3841	CAAAGCGATA	GAAATCATGC	ATGAGCAGAT	CTGCCCCCTG	AGTTTTTCAGG	CAATTTGTGG	AAGCTTGGCG	TCAAGCCTTA
3921	GCGGTAAACG	CAGTGGGACT	CCAACAAGCA	CAATGGTTCA	ATCAGGAAGA	AATGAAAGCC	AACTACAGGG	AAATGGCGAA
4001	GGAGCTTTCT	TGTGGGGACT	TGCTTTGTCA	TTTGCAAACCT	GAGGAGAAGA	CGAGCGTCTT	ACCGAATTCC	CTTCACATCT
4081	TCAACGCCAT	CAAGGGGAAG	GGGAGAGAAA	GGAAATAAAG	CGGGATGACC	AGCTCGTCTT	CGGTGCTGTG	ATTACATCTC
4161	ATGGCCCGTG	CAAGGGGAAG	GGGAGAGAAA	GGAAATAAAG	CAGGATGCTT	TCCAAAGCCA	ATCACTGGGG	AGACCGAGCA
4241	CAGGGAGGAC	CATATTTTTT	TAAATCTCAC	TGGCAATATT	AACAACGTTA	TTTCTTAAAC	GACTTTCTAT	AGGAGTTGTA
4321	AGAAGGTGCA	GACTTAGATT	TTATTTCTCC	TTGCAGAGTA	CAAAGTTTTT	ATTGTGTCTT	AACAAAGGTG	TGGTAGACAC
4401	TCTTGAGCTG	CAGGGAGGGC	TGCACTGACA	TTGATGCCTG	GTGTTAGAAT	AGATGGCCTA	CAGAAAAAAA	AGGTTCTGGG
4481	ATCTACATGG	CAGAACCTAC	TAGTTTTGTC	TAGGAGTATG	GGGGACCTTT	TGCCCTGACT	CGTGCCGGAA	ATCTGATGCT
4561	AATCAGGGTA	AATAGTGAGC	TAACTGAGCT	AGACACTCAA	TTGTATGACT	AGGATTTGTG	CTATTATCTC	ATTCAACAAC
4641	ATAGAGCAAG	CGGGACGGTC	ATATATGTAT	TACATTTCTA	TTAATCCGCT	ACTGGCTTCA	AGTCAGAACT	TTGTCAITAA
4721	TCATCGACTC	TTGTGCTGGT	CCAGTATATG	CAATACACTT	CATTTTAAAT	ACTCACATGG	GCTTATGCAT	TAAGTTTAAAT
4801	TGTGATAAAT	TCCTTACT			TAATGGTTTA	TTCTTGTCTAT	AAAAATGTGC	AATATGGAGA
4881	TGTATACAAG							

FIG. 10D

	10	20	30	40	50	60	70	80
MEGHVMI AFL	PTILNQLFRV	LTRATQEEVA	VNVTRV I HV	VAQCHEEGLE	SHLRSYVKYA	YKAEFYVASE	YKTVHEELTK	80
SMTTILKPSA	DFLT SNKLLR	YSWFFFDVLI	KSMAOHLIEN	SKVKLLRNQR	FPASYHHA AE	TVVNMLMPHI	TOKFGDNPEA	160
SKNANHSLAV	FIKRCPTFMD	RGFVFKQINN	YISCFAPGDP	KTLFEYKFEP	LRVVCNHEHY	IPLNLPMPFG	KGRIQRYQDL	240
QLDYSLTDEF	CRNHFLVGLL	LREVGTAQOE	FREVRLIAIS	VLKNLLIKHS	FDDRYASRSH	QARIATLYLP	LPGLLIENVQ	320
RINVRDVSPF	PVNAGMTVKD	ESLALPAVNP	LVTPOKGSTL	DNSLHDDLLG	AISGLASPYT	TSTPNINSVR	NADSRGSLIS	400
1 TDSGNSLPER	NSEKSNSLDK	HQQSSTLGNS	VVRCDKLDQS	EIKSLIMCFL	YILKMSDDA	LFTYWNKAST	SEIMDFFTIS	480
1 EVCLHQFOYM	GKRYIARNQE	GLGPIVHDRK	SQTLFVSRNR	TGMMHARLOQ	LGSLDNSLTF	NHSYGHSDAD	VLHQSLLEAN	560
1 IATEVCLTAL	DTLSLFTLAF	KNOLLADHGH	NPLMKKVFDV	YLCFLOKHQS	ETALKNVFTA	LRSLIYKFPS	TFYEGRADMC	640
1 AALCYEILKC	CNSKLSSIRT	EASQLLYFLM	RNNFDYTGKK	SFVRTHLOVI	ISVSQLIADV	VGIGETRPOQ	SLSIINN CAN	720
1 SDRLIKHTSF	SSDVKDLTRK	IRTVMATAQ	MKEHENDPEM	LVDLOYSLAK	SYASTPELRA	TWLD SMARIH	VKNGDLSEAA	800
1 MCYVHV TALV	AEYLTRKGVF	ROGCTAFRVI	TPNIDEEASM	MEDVGMQDVH	FNEDVLMELL	EQCADGLWKA	ERYELIADIY	880
1 KLIPIY EKR	RDFFEDEDGK	EYTYKEPKLT	PLSEISQRLI	KLYSDKFGSE	NVIMIQDSGK	VNPKOLD SKY	AYIQVTHVIP	960
51 FFDEKELQER	KTEFERSHNI	RRFMFEMPPT	QTGKROGGVE	EQCKRTILT	AIHCFPYVKK	RIPV MYQHHT	DLNPIEVAID	1040
41 EMSKKVAELR	QLC SSAEVD M	IKLQ LKQGS	VSVQVNAGPL	AYARAFLDDT	NTKRYPDNKV	KLLKEVFRQF	VEACQALAV	1120
21 NERLIKEDQL	EYQEE MKANY	REMAKELSEI	MHEQICPLEE	KTSVLPNSLH	IFNAISGTPT	STMVHGMTSS	SSVV	1194

095833 101300

FIG. 10D (cont.)



10	20	30	40	50	60	70	80
AATTGTAATA	CGACTCACTA	TAGGGCGAAT	TGGGTACCGG	CCCCCCCCCTC	GAGGTGACG	GTATCGATAA	GCTTGATATC 80
GAATTTCGGCA	CGAGTTTAC	ACCATCACCA	AAACCCAGAA	TTTTATGATG	AGATTAAAT	AGAGTTGCC	ACTCAGCTGC 160
ATGAAAAGCA	CCACCTGTTG	CTCACATTCT	TCCATGTCAG	CTGTGACAAC	TCAAGTAAAG	GAAGCACGAA	GAAGAGGGAT 240
GTCGTTGAAA	CCCAAGTTGG	CTACTCCTGG	CTTCCCTCC	TGAAAGACGG	AAGGGTGGTG	ACAAGCGAGC	AGCACATCCC 320
GGTCTCGGCG	AACCTTCCTT	CGGGCTATCT	TGGCTACCAA	GAGCTTGGGA	TGGGCAGGCA	TTATGGTCCG	GAAATTAAAT 400
GGGTAGATGG	AGGCAAGCCA	CTGCTGAAAA	TTTCCACTCA	TCTGGTTTCT	ACAGGGATAC	TCAGGATCAG	CATTTACATA 480
ATTTTTTCCA	GTACTGTCAG	AAAACCGAAT	CTGGAGCCCA	AGCCTTAGGA	AACGAACTTG	TAAAGTACCT	TAAGAGTCTG 560
CATGCGATGG	AAGGCCACGT	GATGATCGCC	TTCTTGCCCA	CTATCCTAAA	CCAGCTGTTT	CGAGTCCCTA	CCAGAGCCAC 640
ACAGGAAGAA	GTCGCGTTA	ACGTGACTCG	GGTCATTATT	CATGTGGTTG	CCCAGTGCCA	TGAGGAAGGA	TTGGAGAGCC 720
ACTTGAGGTC	ATATGTTAAG	TACGCGTATA	AGGCTGAGCC	ATATGTTGCC	TCTGAATACA	AGACAGTGCA	TGAAGAACTG 800
ACCAAATCCA	TGACCACGAT	TCTCAAGCCT	TCTGCCGATT	TCCTCACCAG	CAACAACTA	CTGAGGTACT	CATGGTTTTT 880
CTTTGATGTA	CTGATCAAAT	CTATGGCTCA	GCATTTGATA	GAGAACTCCA	AAGTTAAGTT	GCTGCGAAAC	CAGAGATTTC 960
CTGCATCCTA	TCATCATGCA	GCGGAAACCG	TTGTAAATAT	GCTGATGCCA	CACATCACTC	AGAAGTTTGG	AGATAATCCA 1040
GAGGCATCTA	AGAACGCGAA	TCATAGCCTT	GCTGTCTTCA	TCAAGAGATG	TTTCACCTTC	ATGGACAGGG	GCTTTGCTTT 1120
CAAGCAGATC	AACAACCTACA	TTAGCTGTTT	TGCTCCTGGA	GACCCAAAGA	CCCTCTTTGA	ATACAAGTTT	GAATTTCTCC 1200
GTGTAGTGTG	CAACCATGAA	CATTATATTC	CGTTGAACTT	ACCAATGCCA	TTTGGAAGAG	GCAGGATTCA	AAGATACCAA 1280
GACCTCCAGC	TTGACTACTC	ATTAACAGAT	GAGTTCTGCA	GAAACCACTT	CTTGGTGGGA	CTGTTACTGA	GGGAGGTGGG 1360
GACAGCCCTC	CAGGAGTTCC	GGGAGGTCCG	TCTGATCGCC	ATCAGTGTGC	TCAAGAACCT	TGGGTCTGCT	CATTCTTTTG 1440
ATGACAGATA	TGCTTCAAGG	AGCCATCAGG	CAAGGATAGC	CACCTCTAC	CTGCCCTCTG	TTGGTCTGCT	GATTGAAAC 1520
TCCAGCGGA	TCAATGTGAG	GGATGTGTCA	CCCTTCCCTG	TGAACGCGGG	CATGACCGTG	AAGGATGAAT	CCCTGGCTCT 1600
ACCAGCTGTG	AATCCGCTGG	TGACGCGCA	GAAGGGAAGC	ACCCTGGACA	ACAGCCTGCA	CAAGGACCTG	CTGGGCGCCA 1680
TCTCCGGCAT	TGCTTCTCCA	TATACAACCT	CAACTCCAAA	CATCAACAGT	GTGAGAAATG	CTGATTCGAG	AGGATCTCTC 1760
ATAAGCACAG	ATTCCGGTAA	CAGCCTTCCA	GAAAGGAATA	GTGAGAAGAG	CAATTCCTTG	GATAAGCACC	AACAAAGTAG 1840
CACATTGGGA	AATTCGCTGG	TTCGCTGTGA	TAAACTTGAC	CAGTCTGAGA	TTAAGAGCCT	ACTGATGTGT	TTCTCTTACA 1920
TCTTAAAGAG	CATGCTGTAT	GATGCTTTGT	TTACATATTG	GAACAAGGCT	TCAACATCTG	AACTTATGGA	TTTTTTTACA 2000
ATATCTGAAG	TCTGCCCTGCA	CCAGTTCCAG	TACATGGGGA	AGCGATACAT	AGCCAGGAAC	CAGGAGGGGT	TGGACCCCAT 2080
AGTTTCATGAT	CGAAAGTCTC	AGACATTGCC	TGTTTCCCGT	AACAGAACAG	GAATGATGCA	TGCCAGATTG	CAGCAGCTGG 2160
GCAGCCTGGA	TAACTCTCTC	ACTTTTAAAC	ACAGCTATGG	CCACTCGGAC	GCAGATGTTT	TGCACCAGTC	ATTACTTGAA 2240
GCCACATTTG	CTACTGAGGT	TTGCCCTGACA	GCTCTGGACA	CGCTTCTCT	ATTTACATTG	GCGTTTAAAG	ACCAGCTCCT 2320
GGCGAACCAT	GGACATAATC	CTCTCATGAA	AAAAGTTTTT	GATGTCTACC	TGTGTTTTCT	TCAAAAACAT	CAGTCTGAAA 2400
CGGCTTTAAA	AAATGTCCTC	ACTGCCTTAA	GGTCCTTAAT	TTATAAGTTT	CCCTCAACAT	TCTATGAAGG	GAGAGCGGAC 2480
ATGAGTGCGG	CTCTGTGTTA	CGAGATTCTC	AAGTGCTGTA	ACTCCAAGCT	GAGCTCCATC	AGGACGGAGG	CCTCCAGCT 2560
GCTCTACTTC	CTGATGAGGA	ACAACCTTGA	TTACACTGGA	AAGAAGTCTT	TTGTCCGGAC	ACATTTGCAA	GTCTCATAT 2640
CTGTACGCA	GCTGATAGCA	GACGTTGTTG	GCAATGGGGA	AACCAGATTC	CAGCAGTCCC	TGTCCATCAT	CAACAACTGT 2720
GCCACACATG	ACCGCTTAT	TAAGCACACC	AGCTTCTCCT	CTGATGTGAA	GGACTTAACC	AAAAGGATAC	GCACGGTCT 2800
AATGCCACC	GCCAGATGA	AGGAGCATGA	GAACGACCCA	GAGATGCTGG	TGGACCTCCA	GTACAGCCTG	GCCAAATCCT 2880
ATGECAGCAC	GCCCGAGCTC	AGGAGACGCT	GGCTCGACAG	CATGGCCAGG	ATCCATGTCA	AAAATGGCGA	TCTCTCAGAG 2960
GCAGCAATGT	GCTATGTCCA	CGTAACAGCC	CTAGTGGCAG	AATATCTCAC	ACGGAAAGGC	GTGTTTAGAC	AAGGATGCAC 3040
CGCCTTCAGG	GTCATTACCC	CAAACATCGA	CGAGGAGGCC	TCCATGATGG	AAGACGTGGG	GATGCAGGAT	GTCCATTTC 3120
ACGAGGATGT	GCTGATGGAG	CTCCTTGAGC	AGTGCGCAGA	TGGACTCTGG	AAAGCCGAGC	GCTACGAGCT	CATCGCGGAC 3200
ATCTACAAAC	TTATCATCCC	CATTTATGAG	AAGCGGAGGG	ATTTCTTTGA	AGATGAAGAT	GGAAAGGAGT	ATATTTACAA 3280
GGAAACCAAA	CTCACACCGC	TGTCGGAAT	TTCTCAGAGA	CTCCTTAAAC	TGTACTCGGA	TAAATTTGGT	TCTGAAAATG 3360
TCAAAATGAT	ACAGGATTCT	GGCAAGGTCA	ACCTTAAGGA	TCTGGATTCT	AGATATGCAT	ACATCCAGGT	GACTCACGTC 3440
ATCCCCTTCT	TTGACGAAAA	AGAGTTGCAA	GAAAGGAAAA	CAGAGTTTGA	GAGATCCCAC	AACATCCGCC	GCTTCATGTT 3520
TGAGATGCCA	TTTACGCAGA	CCGGGAAGAG	GCAGGGGGGG	GTGGAAGAGC	AGTGCAAAAC	GOGCACCATC	CTGACAGCCA 3600
TAGACTGCTT	CCCTTATGTG	AAGAAGCGCA	TCCCTGTCTAT	GTACCAGCAC	CACACTGACC	TGAACCCCAT	CGAGGTGGCC 3680
ATTGACGAGA	TGAGTAAGAA	GGTGGCGGAG	CTCCGGCAGC	TGTGCTCCTC	GGCCGAGGTG	GACATGATCA	AACTGCAGCT 3760
CAACTCCAG	GGCAGCGTGA	GTGTTCAAGT	CAATGCTGGC	CCACTAGCAT	ATGCGCGAGC	TTTCTTAGAT	GATACAAACA 3840
CAAAGCGATA	TCCTGACAAT	AAAGTGAAGC	TGCTTAAGGA	AGTTTTTCAGG	CAATTTGTGG	AAGCTTGCGG	TCAAGCCTTA 3920
GCGGTAAACG	AACGTCTGAT	TAAAGAAGAC	CAGCTCGAGT	ATCAGGAAGA	AATGAAAGCC	AACTACAGGG	AAATGGCGAA 4000
GGAGCTTTCT	GAAATCATGC	ATGAGCAGAT	CTGCCCCCTG	GAGGAGAAGA	CGAGCGTCTT	ACCGAATTCC	CTTCACATCT 4080
TCAACGCCAT	CAGTGGGACT	CCAACAAGCA	CAATGGTTCA	CGGGATGACC	AGCTCGTCTT	CGGTCGTGTG	ATTACATCTC 4160
ATGGCCCGTG	TGTGGGGACT	TGCTTTGTCA	TTTGCAAACT	CAGGATGCTT	TCCAAAGCCA	ATCACTGGGG	AGACCGAGCA 4240
CAGGGAGGAC	CAAGGGGAAG	GGGAGAGAAA	GGAAATAAAG	AACAACGTTA	TTTCTTAAAC	GACTTTCTAT	AGGAGTTGTA 4320
AGAAGGTGCA	CATATTTTTT	TAAATCTCAC	TGGCAATATT	CAAAGTTTTT	ATTGTGTCTT	AACAAAGGTG	TGGTAGACAC 4400
TCTTGAGCTG	GACTTAGATT	TTATTCTTCC	TGTCAGAGTA	GTGTTAGAAAT	AGATGGCCTA	CAGAAAAAAA	AGGTTCTGGG 4480
ATCTACATGG	CAGGGAGGGC	TGCACTGACA	TGATGCTTG	GGGGACCTTT	TGCCCTGACT	CGTGCCGGAA	ATCTGATCGT 4560
AATCAGGGTA	CAGAACTTAC	TAGTTTTGTC	TAGGAGTATG	TTGTATGACT	AGGATTTGTG	CTATTATCTC	ATTCAACAAC 4640
ATAGAGCAAG	AATAGTGAGC	TAACTGAGCT	AGACACTCAA	TTAATCCGCT	ACTGGCTTCA	AGTCAGAACT	TTGTCAATTA 4720
TCATCGACTC	CGGGACGGTC	ATATATGTAT	TACATTTCTA	CATTTTTAAT	ACTCACATGG	GCTTATGCAT	TAAGTTTAAT 4800
TGTGATAAAT	TTGTGCTGGT	CCAGTATATG	CAATACACTT	TAATGGTTTA	TTCTTGTCAT	AAAAATGTGC	AATATGGAGA 4880
TGTATACAG	TCTTTACT						4898
10	20	30	40	50	60	70	80

FIG. 10E



10	20	30	40	50	60	70	80
HVMIAFL	PTILNQLFRV	LTRATQEEVA	VNVTRVLIHV	VAOCHEEGLE	SHLRSYVRYA	YKAEFYVASE	YKTVHEELTK 80
TILKPSA	DFTLSNKLRL	YSWFFFDVLI	KSMAQHLEIN	SKVKLLRNQR	FPASYHHAEE	TVVIMLMPHI	TOKFGDNPEA 160
IANHSLAV	FIKRCFTFMD	RGFVFKQINN	YISCFAPGDP	KILFEYKFEF	LRVVCNHEHY	IPLNLPMPFG	KGRIQRYODL 240
YSLTDEF	CRNHFLVGLL	LREVGTAQOE	FREVRLIAIS	VLKNLLIKHS	FDDRYASRSH	QARIATLYLP	LPGLLIENVQ 320
VRDVSPF	PVNAGMTVKD	ESLALPAVNP	LVTPOKGSTL	DNSLHKDLIG	AISGLASPYT	TSTPNINSVR	NADSRGSLIS 400
SGNSLPER	NSEKSNSLDK	HOOSSTLGNS	VVRCDKLDQS	EIKSLIMCFL	YILKSMDDA	LFTYWNKAST	SELMDFFTIS 480
ELHOFQYM	GKRYIARNQE	GLGPIVHDRK	SQTLPVSRNR	TGMMHARLQQ	LGSLDNSLTF	NHSYGHSDAD	VLHQSLLLEAN 560
TEVCLTAL	DTLSLFTLAF	KNOLLADHGH	NPLMKRVFDV	YLCFLOKHQS	ETALKQVPTA	LRSLTYKPPS	TFYEGRADMC 640
LCYEILKC	CNSKLSIRT	EASQLLYFLM	RNNFDYTGKK	SFVRTHLQVI	ISVSQLIADV	VGIGETRPOQ	SLSIINNAN 720
RLIKHTSF	SSDVKDLTKR	IRTVLMATAQ	MKEHENDPEM	LVDLOYSLAK	SYASTPELRK	TWLDSEARH	VINGDLSEAA 800
YVHTALV	AEYLTRKGVF	ROGCTAFRVI	TPNIDEEASM	MEDVGMQDVH	FNEDVLMELL	EQCADGLWKA	ERYELLADIY 880
JIPYIEKR	RDFFEDEDGK	EYTYKEPKLT	PLSEISQRL	KLYSDKFGSE	NVKMIQDSGK	VNPKDLDISKY	AYIQVTHVIP 960
DEKELQER	KTEFERSHNI	RRFMFEMPFT	QTGKROGGVE	EOCKRRTILT	AIHCFFPYVK	RIPVMYQHHT	DLNPIEVAID 1040
4SKKVAELR	OLCSEAEDVM	IKLOLKLOGS	VSVQVAGPL	AYARAFDDT	NTKRYPDNKV	KLLKEVFRQP	VEACQALAV 1120
ERLIKEDQL	EYQEMKANY	REMAKELSEI	MHEQICPLEE	KTSVLPNSLH	IFNAISGTPT	STMVHGMTSS	SSV 1194

09687837-101300

FIG. 10E (cont.)

10	20	30	40	50	60	70	80
AATTGTAATA	CGACTCACTA	TAGGGCGAAT	TGGGTACCGG	GCCCCCCCCTC	GAGGTGACG	GTATCGATAA	GCTTGATATC 80
GAATTCGGCA	CGAGTTTAC	ACCATCACCA	AAACCCAGAA	TTTTATGATG	AGATTAAAT	AGAGTTGCCC	ACTCAGCTGC 160
ATGAAAAGCA	CCACCTGTTG	CTCACATTCT	TCCATGTCAG	CTGTGACAAC	TCAAGTAAAG	GAAGCACGAA	GAAGAGGGAT 240
GTCGTTGAAA	CCCAAGTTGG	CTACTCCTGG	CTTCCCCCTCC	TGAAAGACGG	AAGGGTGGTG	ACAAGCGAGC	AGCACATCCC 320
GGTCTCGGCG	AACCTTCCTT	CGGGCTATCT	TGGCTACCAA	GAGCTTGGGA	TGGGCAGGCA	TTATGGTCCG	GAAATTAAAT 400
GGGTAGATGG	AGGCAAGCCA	CTGCTGAAAA	TTTCCACTCA	TCTGGTTTCT	ACAGGGATAC	TCAGGATCAG	CATTACATA 480
ATTTTTTCCA	GTAATGTCAG	AAAACCGAAT	CTGGAGCCCC	AGCCTTAGGA	AACGAACCTG	TAAAGTACCT	TAAGAGTCTG 560
CATGCCATGG	AAGGCCACGT	GATGATCGCC	TTCTTGCCCA	CTATCCTAAA	CCAGCTGTTT	CGAGTCCTCA	CCAGAGCCAC 640
ACAGGAAGAA	GTCGCGGTTA	ACGTGACTCG	GGTCATTATT	CATGTGGTTG	CCCAGTGCCA	TGAGGAAGGA	TTGGAGAGCC 720
ACTTGAGGTC	ATATGTTAAG	TACGCGTATA	AGGCTGAGCC	ATATGTTGCC	TCTGAATACA	AGACAGTGCA	TGAAGAAGTG 800
ACCAAATCCA	TGACCACGAT	TCTCAAGCCT	TCTGCCGATT	TCCTCACCAG	CAACAACTA	CTGAGGTACT	CATGGTTTTT 880
CTTTGATGTA	CTGATCAAAT	CTATGGCTCA	GCATTTGATA	GAGAACTCCA	AAGTTAAGTT	GCTGCGAAAC	CAGAGATTTT 960
CTGCATCCTA	TCATCATGCA	GCGGAAACCG	TTGTAATAT	GCTGATGCCA	CACATCACTC	AGAAGTTTGG	AGATAATCCA 1040
GAGGCATCTA	AGAACGCGAA	TCATAGCCTT	GCTGTCTTCA	TCAAGAGATG	TTTACCTTTC	ATGGACAGGG	GCTTTGTCTT 1120
CAAGCAGATC	AACAACCTACA	TTAGCTGTTT	TGCTCCTGGA	GACCCAAAGA	CCCTCTTTGA	ATACAAGTTT	GAATTTCTCC 1200
GTGTAGTGTG	CAACCATGAA	CATTATATTC	CGTTGAACCT	ACCAATGCCA	TTTGGAAAAG	GCAGGATTCA	AAGATACCAA 1280
GACCTCCAGC	TTGACTACTC	ATTAACAGAT	GAGTTCTGCA	GAAACCACTT	CTTGGTGGGA	CTGTTACTGA	GGGAGGTGGG 1360
GACAGCCCTC	CAGGAGTTCC	GGGAGGTCCG	TCTGATCGCC	ATCAGTGTGC	TCAAGAACCT	GCTGATAAAG	CATTCTTTTG 1440
ATGACAGATA	TGCTTCAAGG	AGCCATCAGG	CAAGGATAGC	CACCTCTAC	CTGCCCTGT	TGGTCTGCT	GATTGAAAAC 1520
GTCCAGCGTA	TCAATGTGAG	GGATGTGTCA	CCCTTCCCTG	TGAAACGCGG	CATGACCGTG	AAGGATGAAT	CCCTGGCTCT 1600
ACCAGCTGTG	AATCCGCTGG	TGACGCGCGA	GAAGGGAAGC	ACCCTGGACA	ACAGCCTGCA	CAAGGACCTG	CTGGGCGCCA 1680
TCTCCGGCAT	TGCTTCTCCA	TATACAACCT	CAACTCCAAA	CATCAACAGT	GTGAGAAATG	CTGATTCGAG	AGGATCTCTC 1760
ATAAGCACAG	ATTCCGGTAA	CAGCCTTCCA	GAAAGGAATA	GTGAGAAGAG	CAATTCCCTG	GATAAGCACC	AACAAAGTAG 1840
CACATTGGGA	AATTCGCTGG	TTCGCTGTGA	TAAACTTGAC	CAGTCTGAGA	TTAAGAGCCT	ACTGATGTGT	TTCTCTTACA 1920
TCTTAAAGAG	CATGTCTGAT	GATGCTTTGT	TTACATATTG	GAACAAGGCT	TCAACATCTG	AACTTATGGA	TTTTTTTACA 2000
ATATCTGAAG	TCTGCCGTGA	CCAGTTCCAG	TACATGGGGA	AGCGATACAT	AGCCAGGAAC	CAGGAGGGGT	TGGGACCCAT 2080
AGTTTCATGAT	CGAAAGTCTC	AGACATTGCC	TGTTTCCCGT	AACAGAACAG	GAATGATGCA	TGCCAGATTG	CAGCAGCTGG 2160
GCAGECTGGA	TAACTCTCTC	ACTTTTAACC	ACAGCTATGG	CCACTCGGAC	GCAGATGTTT	TGCACCAGTC	ATTACTTGAA 2240
GCCACATTTG	CTACTGAGGT	TTGCCGTGCA	GCTCTGGACA	CGCTTTCTCT	ATTTACATTG	GCGTTTAAGA	ACCAGCTCCT 2320
GGCEGACCAT	GGACATAATC	CTCTCATGAA	AAAAGTTTTT	GATGTCTACC	TGTGTTTTCT	TCAAAAACAT	CAGTCTGAAA 2400
CGGCTTTAAA	AAATGTCTTC	ACTGCCTTAA	GGTCCTTAAT	TTATAAGTTT	CCCTCAACAT	TCTATGAAGG	GAGAGCGGAC 2480
ATGATGCGGG	CTCTGTGTTA	CGAGATTCTC	AAGTGCTGTA	ACTCCAAGCT	GAGCTCCATC	AGGAACGGAGG	CCTCCAGCT 2560
GCTCTACTTC	CTGATGAGGA	ACAACCTTGA	TTACACTGGA	AAGAAGTCTT	TTGTCCGGAC	ACATTTGCAA	GTCTCATAT 2640
CTGTGAGCCA	GCTGATAGCA	GACGTTGTTG	GCATTGGGGA	AACCAGATTG	CAGCAGTCCC	TGTCCATCAT	CAACAACGTG 2720
GCCAACAGTG	ACCGGCTTAT	TAAGCACACC	AGCTTCTCCT	CTGATGTGAA	GGACTTAACC	AAAAGGATAC	GCACGGTGCT 2800
AATGGCCACC	CCCCAGATGA	AGGAGCATGA	GAACGACCCA	GAGATGCTGG	TGGACCTCCA	GTACAGCCTG	GCCAAATCCT 2880
ATGGEAGCAC	GCCCGAGCTC	AGGAAGACGT	GGCTCGACAG	CATGGCCAGG	ATCCATGTCA	AAAATGGCGA	TCTCTCAGAG 2960
GCAGCAATGT	GCTATGTCCA	CGTAACAGCC	CTAGTGCCAG	AATATCTCAC	ACGGAAAGGC	GTGTTTAGAC	AAGGATGCAC 3040
CGCCTTCAGG	GTCAATTACC	CAAACATGSA	CGAGGAGGCC	TCCATGATGG	AAGACGTGGG	GATGCAGGAT	GTCCATTTC 3120
ACGAGGATGT	GCTGATGGAG	CTCCTTGAGC	AGTGCCGAGA	TGGACTCTGG	AAAGCCGAGC	GCTACGAGCT	CATCGCCGAC 3200
ATCTACAAAC	TTATCATCCC	CATTATGAG	AAGCGGAGGG	ATTTCTTTGA	AGATGAAGAT	GGAAAGGAGT	ATATTTACAA 3280
GGAACTCCAA	CTCACACCGC	TGTGGAAT	TTCTCAGAGA	CTCCTTAAAC	TGTACTCGGA	TAAATTTGGT	TCTGAAAATG 3360
TCAAAATGAT	ACAGGATTCT	GGCAAGGTCA	ACCTTAAGGA	TCTGGATTCT	AAGTATGCAT	ACATCCAGGT	GACTCACGTC 3440
ATGECCTTCT	TTGACGAAAA	AGAGTTGCAA	GAAAGGAAAA	CAGAGTTTGA	GAGATCCCAC	AACATCCGCG	GCTTCATGTT 3520
TGAGATGCCA	TTTACGCAGA	CCGGGAAGAG	GCAGGGCGGG	GTGGAAGAGC	AGTGCAAACG	GCGCACCATC	CTGACAGCCA 3600
TACACTGCTT	CCCTTATGTG	AAGAAGCGCA	TCCCTGTCTAT	GTACCAGCAC	CACACTGACC	TGAACCCCAT	CGAGGTGGCC 3680
ATPGACGAGA	TGAGTAAGAA	GGTGGCGGAG	CTCCGGCAGC	TGTGCTCCTC	GGCCGAGGTG	GACATGATCA	AACTGCAGCT 3760
CAAACTCCAG	GGCAGCGTGA	GTGTTCAAGT	CAATGCTGGC	CCACTAGCAT	ATGCCGAGC	TTTCTTAGAT	GATACAAACA 3840
CAAAGCGATA	TCCTGACAA	AAAGTGAAGC	TGCTTAAGGA	AGTTTTTCA	CAATTTGTTG	AAGCTTGCGG	TCAAGCCTTA 3920
GCGGTAAACG	AACGTCTGAT	TAAAGAAGAC	CAGCTCGAGT	ATCAGGAAGA	AATGAAAGCC	AACTACAGGG	AAATGGCGAA 4000
GGAGCTTTCT	GAAATCATGC	ATGAGCAGAT	CTGCCCCCTG	GAGGAGAAGA	CGAGCGTCTT	ACCGAATTCC	CTTCACATCT 4080
TCAACGCCAT	CAGTGGGACT	CCAACAAGCA	CAATGGTTCA	CGGGATGACC	AGCTCGTCTT	CGGTCTGTGT	ATTACATCTC 4160
ATGGCCCGTG	TGTGGGGACT	TGCTTTGTCA	TTTGCAAAC	CAGGATGCTT	TCCAAAGCCA	ATCACTGGGG	AGACCGAGCA 4240
CAGGGAGGAC	CAAGGGGAAG	GGGAGAGAAA	GGAAATAAAG	AACAACGTTA	TTTCTTAACA	GACTTTCTAT	AGGAGTTGTA 4320
AGAAGGTGCA	CATATTTTTT	TAAATCTCAC	TGGCAATATT	CAAAGTTTTT	ATTGTGTCTT	AACAAAGGTG	TGGTAGACAC 4400
TCTTGAGCTG	GACTTAGATT	TTATTCTTCC	TTGCAGAGTA	GTGTTAGAAT	AGATGGCCTA	CAGAAAAAAA	AGGTTCTGGG 4480
ATCTACATGG	CAGGAGGGGC	TGCATGACA	TTGATGCCCTG	GGGGACCTTT	TGCCCTGACT	CGTGCCGGAA	ATCTGATCGT 4560
AATCAGGGTA	CAGAACTTAC	TAGTTTTGTC	TAGGAGTATG	TTGTATGACT	AGGATTTGTG	CTATTATCTC	ATTCAACAAC 4640
ATAGAGCAAG	AATAGTGAGC	TAACTGAGCT	AGACACTCAA	TTAATCGCT	ACTGGCTTCA	AGTCAGAACT	TTGTCAATTA 4720
TCATCGACTC	CGGGACGGTC	ATATATGTAT	TACATTTCTA	CATTTTAAAT	ACTCACATGG	GCTTATGCAT	TAAGTTTAA 4800
TGTGATAAAT	TTGTGCTGGT	CCAGTATATG	CAATACACTT	TAATGGTTTA	TTCTTGTCAT	AAAAATGTGC	AATATGGAGA 4880
TGTATACAAG	TCITTAAT						4898
10	20	30	40	50	60	70	80

FIG. 10F

10	20	30	40	50	60	70	80
HVMIAFL	PTILNQLFRV	LTRATQEEVA	VNVTRVHHV	VAQCHEEGLE	SHLSYVKA	YKAEPYVASE	YKTVHEELTK 80
TILKPSA	DFLTSNKLLR	YSWFFFDVLI	KSMAQHLEN	SKVKLLRNOR	FPASYHAAE	TVVMMLMPHI	TOKFGDNFEA 160
ANHSLAV	FIKRCFTFMD	RGFVFKQDNN	YISCFAPGDP	KTLFEYKPEF	LRVVCNMEHY	IPLNLMPFFG	KGRIQRYODL 240
YSLTDEF	CRNHFLVGLL	LREVGTLAQE	FREVRLLAIS	VLNLLIKHS	FDDRYASRSH	QARIATLYLP	LPGILLIENVQ 320
NRDVSPF	PVNAGMTVKD	ESLALPAVNP	LVTPOKGSTL	DNSLHKDLIG	AISGLASPYT	TSTFNINSVR	NADSRGSLIS 400
GNLSLPER	NSEKSNLSDK	HQOSSTLGNS	VVRCDKLDQS	EIKSLIMCFL	YILKSMDDA	LFTYWNKAST	SELMOFFTIS 480
LLHQFOYM	GKRYIARNOE	GLGPIVHDKK	SQTLFVERNR	TGMMHARLOQ	LGSLDNSLTF	NHSYGHSDAD	VLHOSLLEAN 560
FEVCLTAL	DTLSLFTLAF	KNOLLADHGH	NPLMKKVFDV	YLCFLOKHQS	ETALJQVFTA	LRSLTYKFPS	TFYEGRADMC 640
UCYEILKC	CNSKLSSTRT	EASQLLYFLM	RNNFDYTGKK	SFVRTHLOVI	ISVSQLIADV	VGIGETRPOQ	SLSTIDNCCAN 720
RLIKHTSF	SSDVKDLTKR	IRTVLMATAQ	MKEHENDEPM	LVDLOYSLAK	SYASTPELRK	TWLDSEARH	VINGDLSEAA 800
YVHVITALV	AEYLTRKGVP	ROGCTAFRVI	TPNIDEEASM	MEDVGMQDVH	FNEQVIMELL	EQCADGLWKA	ERYELIADTY 880
IPIIYEKR	RDFFEDEDGK	EYIYKEPKLT	PLSEISQRL	KLYSDKFGSE	NVMIQDSGK	VNPKDLDSKY	AYIQVTHVIP 960
DEKELQER	KTEFERSHNI	RRFMFEMPFT	QTGKROGGVE	EOCKRRTILT	AIHCFFPVKK	RIFVMYQHHT	DLNPIEVAID 1040
ISKVAEALR	QLCSSAEVDM	IKLQWKLOGS	VSVQVNAQPL	AYARAFIDDT	NTRYPDONKV	KLLKEVFROP	VEACQALAV 1120
RLIKEDQL	EYQDEMKANY	REMAKELSEI	MHEQICPLEE	KTSVLPNSLH	IFNAISGTPT	STMVHGMTSS	SSVV 1194

FIG. 10F (cont.)

	10	20	30	40	50	60	70	80	
1	AATTGTAATA	CGACTCACTA	TAGGGCGAAT	TGGGTACCGG	GCCCCCCTC	GAGGTGACG	GTATCGATAA	GCTTGATATC	80
81	GAATTCGGCA	CGAGTTTTAC	ACCATCACCA	AAACCCAGAA	TTTTATGATG	AGATTAAAAT	AGAGTTGCC	ACTCAGCTGC	160
161	ATGAAAAGCA	CCACCTGTTG	CTCACATTCT	TCCATGTCAG	CTGTGACAAC	TCAAGTAAAG	GAAGCACGAA	GAAGAGGGAT	240
241	GTCGTTGAAA	CCCAAGTTGG	CTACTCCTGG	CTTCCCCTCC	TGAAAGACGG	AAGGGTGGTG	ACAAGCGAGC	AGCACATCCC	320
321	GGTCTCGGCG	AACCTTCCTT	CGGGCTATCT	TGGCTACCAA	GAGCTTGGGA	TGGGCAGGCA	TTATGGTCCG	GAAATTAAAT	400
401	GGGTAGATGG	AGGCAAGCCA	CTGCTGAAAA	TTTCCACTCA	TCTGGTTTCT	ACAGGGATAC	TCAGGATCAG	CATTTACATA	480
481	ATTTTTTCCA	GTACTGTCAG	AAAACCGAAT	CTGGAGCCCA	AGCCTTAGGA	AACGAACTTG	TAAAGTACCT	TAAGAGTCTG	560
561	CATGCGATGG	AAGGCCACGT	GATGATCGCC	TTCTTGCCCA	CTATCCTAAA	CCAGCTGTTC	CGAGTCTCTA	CCAGAGCCAC	640
641	ACAGGAAGAA	GTCGCGGTTA	ACGTGACTCG	GGTCATTATT	CATGTGGTTG	CCCAGTGCCA	TGAGGAAGGA	TTGGAGAGCC	720
721	ACTTGAGGTC	ATATGTTAAG	TACGCGTATA	AGGCTGAGCC	ATATGTTGCC	TCTGAATACA	AGACAGTGCA	TGAAGAAGTG	800
801	ACCAAATCCA	TGACCACGAT	TCTCAAGCCT	TCTGCCGATT	TCTCACCAG	CAACAACTA	CTGAGGTA	CATGGTTTTT	880
881	CTTTGATGTA	CTGATCAAAT	CTATGGCTCA	GCATTTGATA	GAGAACTCCA	AAGTTAAGTT	GCTGCGAAAC	CAGAGATTTT	960
961	CTGCATCCTA	TCATCATGCA	GCGGAAACCG	TTGTAATAT	GCTGATGCCA	CACATCACTC	AGAAGTTTGG	AGATAATCCA	1040
1041	GAGGCATCTA	AGAACGCGAA	TCATAGCCTT	GCTGTCTTCA	TCAAGAGATG	TTTCACCTTC	ATGGACAGGG	GCTTTGTCTT	1120
1121	CAAGCAGATC	AACAACCTACA	TTAGCTGTTT	TGCTCCTGGA	GACCCAAAGA	CCCTCTTTGA	ATACAAGTTT	GAATTTCTCC	1200
1201	GTGTAGTGTG	CAACCATGAA	CATTATATTC	CGTTGAACTT	ACCAATGCCA	TTTGGAAAAG	GCAGGATTCA	AAGATACCAA	1280
1281	GACCTCCAGC	TTGACTACTC	ATTAACAGAT	GAGTTCTGCA	GAAACCACTT	CTTGGTGGGA	CTGTTACTGA	GGGAGGTGGG	1360
1361	GACAGCCCTC	CAGGAGTTCC	GGGAGGTCCG	TCTGATCGCC	ATCAGTGTGC	TCAAGAACCT	GCTGATAAAG	CATTCCTTTT	1440
1441	ATGACAGATA	TGCTTCAAGG	AGCCATCAGG	CAAGGATAGC	CACCTCTCTG	CTGCCCTCTG	TTGGTCTGCT	GATTGAAAAC	1520
1521	GTCCAGCGGA	TCAATGTGAG	GGATGTGTCA	CCCTTCCCTG	TGAACGCGGG	CATGACCGTG	AAGGATGAAT	CCCTGGCTCT	1600
1601	ACCAGCTGTG	AATCCGCTGG	TGACGCGGCA	GAAGGGAAGC	ACCTTGGA	ACAGCCTGCA	CAAGGACCTG	CTGGGCGCCA	1680
1681	TCTCCGGCAT	TGCTTCTCCA	TATACAACCT	CAACTCCAAA	CATCAACAGT	GTGAGAAATG	CTGATTGAG	AGGATCTCTC	1760
1761	ATAAGCACAG	ATTCGGGTAA	CAGCCTTCCA	GAAAGGAATA	GTGAGAAGAG	CAATTCCCTG	GATAAGCACC	AACAAAGTAG	1840
1841	CATTTGGGA	AATTCCGTGG	TTGCTGTGA	TAAACTTGAC	CAGTCTGAGA	TTAAGAGCCT	ACTGATGTGT	TTCTCTTACA	1920
1921	TCTTAAAGAG	CATGCTGTGAT	GATGCTTTGT	TTACATATTG	GAACAAGGCT	TCAACATCTG	AACTTATGGA	TTTTTTTACA	2000
2001	ATATCTGAAG	TCTGCTGCA	CCAGTTCCAG	TACATGGGGA	AGCGATACAT	AGCCAGGAAC	CAGGAGGGGT	TGGGACCCAT	2080
2081	AGTTCATGAT	CGAAAGTCTC	AGCATTTGCC	TGTTTCCOGT	AACAGAACAG	GAATGATGCA	TGCCAGATTG	CAGCAGCTGG	2160
2161	GCAGCCTGGA	TAACCTCTCTC	ACTTTTAAAC	ACAGCTATGG	CCACTCGGAC	GCAGATGTTT	TGCACCAGTC	ATTACTTGAA	2240
2241	GCCAACATTG	CTACTGAGGT	TTGCTTGACA	GCTCTGGACA	CGCTTCTCTT	ATTTACATTG	GCGTTTAAAG	ACCAGCTCCT	2320
2321	GGCCGACCAT	GGACATAATC	CTCTCATGAA	AAAAGTTT	GATGCTTACC	TGTGTTTTCT	TCAAAAACAT	CAGTCTGAAA	2400
2401	CGGCTTTAAA	AAATGTCTTC	ACTGCCTTAA	GGTCTTAAAT	TTATAAGTTT	CCCTCAACAT	TCTATGAAGG	GAGAGCGGAC	2480
2481	ATGTGTGCGG	CTCTGTGTTA	CGAGATTCTC	AAGTGCTGTA	ACTCCAAGCT	TTGTCCGAGC	ACATTTGCAA	GTCATCATAT	2560
2561	GCTCTACTTC	CTGATGAGGA	ACAACCTTGA	TTACACTGGA	AAGAAGTCTT	TTGTCCGAGC	ACATTTGCAA	GTCATCATAT	2640
2641	CTGTGAGCCA	GCTGATAGCA	GACGTTGTTG	GCATTTGGGA	AACCAGATTG	CAGCAGTCCC	TGTCCATCAT	CAACAAGTGT	2720
2721	GCCAACAGTG	ACCGGCTTAT	TAAGCACACC	AGCTTCTCCT	CTGATGTGAA	GGACTTAAAC	AAAAGGATAC	GCACGGTCT	2800
2801	AATGGCCACC	GCCAGATGA	AGGAGCATGA	GAACGACCCA	GAGATGCTGG	TGGACCTCCA	GTACAGCCTG	GCCAAATCCT	2880
2881	ATGCCAGCAC	GCCCGAGCTC	AGGAAGACGT	GGCTCGACAG	CATGGCCAGG	ATCCATGTCA	AAAATGGCGA	TCTCTCAGAG	2960
2961	GCAGCAATGT	GCTATGTCCA	CGTAACAGCC	CTAGTGGCAG	AATATCTCAC	ACCGAAAGGC	GTGTTTAGAC	AAGGATGCAC	3040
3041	CGCCTTCAGG	GTCATTACCC	CAACATCGA	CGAGGAGGCC	TCCATGATGG	AAGACGTGGG	GATGCAGGAT	GTCCATTTC	3120
3121	ACGAGGATGT	GCTGATGGAG	CTCCTTGAGC	AGTGGCAGA	TGGACTCTGG	AAAGCCGAGC	GCTACGAGCT	CATCGCCGAC	3200
3201	ATCTACAAAC	TTATCATCCC	CATTTATGAG	AAGCGGAGGG	ATTTCTTTGA	AGATGAAGAT	GGAAAGGAGT	ATATTTACAA	3280
3281	GGAACCCAAA	CTCACACCGC	TGTGCGAAAT	TTCTCAGAGA	CTCCTTAAAC	TGTACTCGGA	TAAATTTGGT	TCTGAAAATG	3360
3361	TCAAAATGAT	ACAGGATTCT	GGCAAGGTCA	ACCCTAAGGA	CTCGGATTCT	AAGTATGCAT	ACATCCAGGT	GACTCACGTC	3440
3441	ATCCCCCTCT	TTGACGAAAA	AGAGTTGCAA	GAAAGGAAAA	CAGAGTTTGA	GAGATCCCAC	AACATCCGCC	GCTTCATGTT	3520
3521	TGAGATGCCA	TTTACGCGAG	CCGGGAAGAG	GCAGGGCGGG	GTGGAAGAGC	AGTGCAAACG	GCGCACCATC	CTGACAGCCA	3600
3601	TACACTGCTT	CCCTTATGTG	AAGAAGCGCA	TCCCTGTCTAT	GTACCAGCAC	CACACTGACC	TGAACCCCAT	CGAGGTGGCC	3680
3681	ATTGACGAGA	TGAGTAAGAA	GGTGGCGGAG	CTCCGGCAGC	TGTGCTCCTC	GGCCGAGGTG	GACATGATCA	AACTGCAGCT	3760
3761	CAAACCTCCAG	GGCAGCGTGA	GTGTTTCAAGT	CAATGCTGGC	CCACTAGCAT	ATGCGCGAGC	TTTCTTAGAT	GATACAAACA	3840
3841	CAAAGCGATA	TCCTGACAAT	AAAGTGAAGC	TGCTTAAGGA	AGTTTTCAGG	CAATTTGTGG	AAGCTTGCGG	TCAAGCCTTA	3920
3921	GCGGTAAACG	AACGTCTGAT	TAAAGAAGAC	CAGCTCGAGT	ATCAGGAAGA	AATGAAAGCC	AACTACAGGG	AAATGGCGAA	4000
4001	GGAGCTTTCT	GAAATCATGC	ATGAGCAGAT	CTGCCCCCTG	GAGGAGAAGA	CGAGCGTCTT	ACCGAATTCC	CTTCACATCT	4080
4081	TCAACGCCAT	CAGTGGGACT	CCAACAAGCA	CAATGGTTCA	CGGGATGACC	AGCTCGTCTT	CGGTCGTGTG	ATTACATCTC	4160
4161	ATGGCCCGTG	TGTGGGGACT	TGCTTTGTCA	TTTGCAAACT	CAGGATGCTT	TCCAAAGCCA	ATCACTGGGG	AGACCGAGCA	4240
4241	CAGGGAGGAC	CAAGGGGAAG	GGGAGAGAAA	GGAATAAAG	AACAACGTTA	TTTCTTAAAC	GACTTCTTAT	AGGAGTTGTA	4320
4321	AGAAGGTGCA	CATATTTTTT	TAAATCTCAC	TGGCAATATT	CAAAGTTTTC	ATTGTGTCTT	AACAAAGGTG	TGGTAGACAC	4400
4401	TCTTGAGCTG	GACTTAGATT	TTATCTTCC	TTGCAGAGTA	GTGTTAGAA	AGATGGCCTA	CAGAAAAAAA	AGGTTCTGGG	4480
4481	ATCTACATGG	CAGGGAGGGC	TGCACTGACA	TTGATGCCTG	GGGGACCTTT	TGCCCTGACT	CGTGCCGGAA	ATCTGATCGT	4560
4561	AATCAGGGTA	CAGAACTTAC	TAGTTTTGTC	TAGGAGTATG	TTGTATGACT	AGGATTTGTG	CTATTATCTC	ATTCAACAAC	4640
4641	ATAGAGCAAG	AATAGTGAGC	TAACTGAGCT	AGACACTCAA	TTAATCCGCT	ACTGGCTTCA	AGTCAGAACT	TTGTCAATTA	4720
4721	TCATCGACTC	CGGGACGGTC	ATATATGTAT	TACATTTCTA	CATTTTAAAT	ACTCACATGG	GCTTATGCAT	TAAGTTTAAAT	4800
4801	TGTGATAAAT	TTGTGCTGGT	CCAGTATATG	CAATACACTT	TAATGGTTTA	TTCTTGTCAT	AAAAATGTGC	AATATGGAGA	4880
4881	TGTATACAAG	TCTTTACT							4898

Fig. 10G





	10	20	30	40	50	60	70	80
1	AATTGTAATA	CGACTCACTA	TAGGGCGAAT	TGGGTACCGG	GCCCCCCTC	GAGGTCGACG	GTATCGATAA	GCTTGATATC 80
81	GAATTCGGCA	CGAGTTTTAC	ACCATCACCA	AAACCCAGAA	TTTTATGATG	AGATTAAAT	AGAGTTGCC	ACTCAGCTGC 160
161	ATGAAAAGCA	CCACCTGTTG	CTCACATTCT	TCCATGTCAG	CTGTGACAAC	TCAAGTAAAG	GAAGCACGAA	GAAGAGGGAT 240
241	GTCGTTGAAA	CCCAAGTTGG	CTACTCCTGG	CTTCCCCTCC	TGAAAGACGG	AAGGGTGGTG	ACAAGCGAGC	AGCACATCCC 320
321	GGTCTCGGCG	AACCTTCCTT	CGGGCTATCT	TGGCTACCAA	GAGCTTGGGA	TGGGCAGGCA	TTATGGTCCG	GAAATTAAAT 400
401	GGGTAGATGG	AGGCAAGCCA	CTGCTGAAAA	TTTCCACTCA	TCTGGTTTCT	ACAGGGATAC	TCAGGATCAG	CATTTACATA 480
481	ATTTTTTCCA	GTACTGTCAG	AAAACCGAAT	CTGGAGCCCA	AGCCTTAGGA	AACGAACCTG	TAAAGTACCT	TAAGAGTCTG 560
561	CATGCGATGG	AAGGCCACGT	GATGATCGCC	TTCTTGCCCA	CTATCCTAAA	CCAGCTGTTC	CGAGTCCTCA	CCAGAGCCAC 640
641	ACAGGAAGAA	GTCGCGGTTA	ACGTGACTCG	GGTCATTATT	CATGTGGTTG	CCCAGTGCCA	TGAGGAAGGA	TTGGAGAGCC 720
721	ACTTGAGGTC	ATATGTTAAG	TACGCGTATA	AGGCTGAGCC	ATATGTTGCC	TCTGAATACA	AGACAGTGCA	TGAAGAACTG 800
801	ACCAAATCCA	TGACCACGAT	TCTCAAGCCT	TCTGCCGATT	TCCTCACCAG	CAACAACTA	CTGAGGTACT	CATGGTTTTT 880
881	CTTTGATGTA	CTGATCAAAT	CTATGGCTCA	GCAATTTGATA	GAGAACTCCA	AAGTTAAGTT	GCTGCGAAAC	CAGAGATTTC 960
961	CTGCATCCTA	TCATCATGCA	GCGGAAACCG	TTGTAAATAT	GCTGATGCCA	CACATCACTC	AGAAGTTTGG	AGATAATCCA 1040
1041	GAGGCATCTA	AGAACGCGAA	TCATAGCCTT	GCTGTCTTCA	TCAAGAGATG	TTTCACCTTC	ATGGACAGGG	GCTTTGTCTT 1120
1121	CAAGCAGATC	AACAACCTACA	TTAGCTGTTT	TGCTCCTGGA	GACCCAAAGA	CCCTCTTTGA	ATACAAGTTT	GAATTTCTCC 1200
1201	GTGTAGTGTG	CAACCATGAA	CATTATATTC	CGTTGAACTT	ACCAATGCCA	TTTGGAAAAG	GCAGGATTCA	AAGATACCAA 1280
1281	GACCTCCAGC	TTGACTACTC	ATTAACAGAT	GAGTTCTGCA	GAAACCACTT	CTTGGTGGGA	CTGTTACTGA	GGGAGGTGGG 1360
1361	GACAGCCCTC	CAGGAGTTCC	GGGAGGTCCG	TCTGATCGCC	ATCAGTGTGC	TCAAGAACCT	GCTGATAAAG	CATTCCTTTG 1440
1441	ATGACAGATA	TGCTTCAAGG	AGCCATCAGG	CAAGGATAGC	CACCTCTAC	CTGCCCTGT	TTGGTCTGCT	GATTGAAAAC 1520
1521	GTCCAGCGGA	TCAATGTGAG	GGATGTGTCA	CCCTTCCCTG	TGAACGCGGG	CATGACCGTG	AAGGATGAAT	CCCTGGCTCT 1600
1601	ACCAGCTGTG	AATCCGCTGG	TGACGCGCA	GAAGGGAAGC	ACCTTGACA	ACAGCCTGCA	CAAGGACCTG	CTGGGCGCCA 1680
1681	TCTCCGGCAT	TGCTTCTCCA	TATACAACCT	CAACTCCAAA	CATCAACAGT	GTGAGAAATG	CTGATTCGAG	AGGATCTCTC 1760
1761	ATAAGCACAG	ATTGCGGTAA	CAGCCTTCCA	GAAAGGAATA	GTGAGAAGAG	CAATTCCCTG	GATAAGCACC	AACAAAGTAG 1840
1841	CACATTGGGA	AATTCCGTGG	TTGCTGTGA	TAAACTTGAC	CAGTCTGAGA	TTAAGAGCCT	ACTGATGTGT	TTCTCTTACA 1920
1921	TCTTAAAGAG	CATGTCTGAT	GATGCTTTGT	TTACATATTG	GAACAAGGCT	TCAACATCTG	AACTTATGGA	TTTTTTTACA 2000
2001	ATATCTGAAG	TCTGCCTGCA	CCAGTTCCAG	TACATGGGGA	AGCGATACAT	AGCCAGGAAC	CAGGAGGGGT	TGGGACCCAT 2080
2081	AGTTCATGAT	CGAAAGTCTC	AGACATTGCC	TGTTTCCCGT	AACAGAACAG	GAATGATGCA	TGCCAGATTG	CAGCAGCTGG 2160
2161	GCAGCCTGGA	TAACCTCTCT	ACTTTTAAAC	ACAGCTATGG	CCACTCGGAC	CGATATGTTT	TGCACCAGTC	ATTACTTGAA 2240
2241	GCCAACATG	CTACTGAGGT	TTGCCTGACA	GCTCTGGACA	CGCTTCTCT	ATTACATTTG	GCGTTTAAAG	ACCAGCTCCT 2320
2321	GGCCGACCAT	GGACATAATC	CTCTCATGAA	AAAAGTTTTT	GATGTCTACC	TGTGTTTTCT	TCAAAAACAT	CAGTCTGAAA 2400
2401	CGGCTTTAAA	AAATGTCCTC	ACTGCCTTAA	GGTCCCTAAT	TTATAAGTTT	CCCTCAACAT	TCTATGAAGG	GAGAGCGGAC 2480
2481	ATGTGTGCGG	CTCTGTGTTA	CGAGATTCTC	AAGTGTGTGA	ACTCCAAGCT	GAGCTCCATC	AGGACGGAGG	CCTCCAGCT 2560
2561	GCTCTACTTC	CTGATGAGGA	ACAACCTTGA	TTACACTGGA	AAGAAGTCCT	TTGTCCGGAC	ACATTTGCAA	GTCATCATAT 2640
2641	CTGTACGCCA	CCTGATAGCA	GACGTTGTGG	GCATTGGGGA	AACCAGATTG	CAGCAGTCCC	TGTCCATCAT	CAACAACGTG 2720
2721	GCCAACAGTG	ACCGGCTTAT	TAAGCACACC	AGCTTCTCCT	CTGATGTGAA	GGACTTAACC	AAAAGGATAC	GCACGGTGCT 2800
2801	AATGGCCACC	GCCCAGATGA	AGGAGCATGA	GAAAGACCCA	GAGATGCTGG	TGGACCTCCA	GTACAGCCTG	GCCAAATCCT 2880
2881	ATGCCAGCAC	GCCCGAGCTC	AGGAAGACGT	GGCTCGACAG	CATGGCCAGG	ATCCATGTCA	AAAATGGCGA	TCTCTCAGAG 2960
2961	GCAGCAATGT	GCTATGTCCA	CGTAACAGCC	CTAGTGGCAG	AATATCTCAC	ACGGAAAGGC	GTGTTTAGAC	AAGGATGCAC 3040
3041	CGCCTTCAGG	GTCATTACCC	CAAACATCGA	CGAGGAGGCC	TCCATGATGG	AAGACGTGGG	GATGCAGGAT	GTCCATTTCA 3120
3121	ACGAGGATGT	GCTGATGGAG	CTCCTTGAGC	AGTGCAGCAG	TGGACTCTGG	AAAGCCGAGC	GCTACGAGCT	CATCGCGGAC 3200
3201	ATCTACAAAC	TTATCATCCC	CATTTATGAG	AAGCGGAGGG	ATTTCTTTGA	AGATGAAGAT	GGAAAGGAGT	ATATTTACAA 3280
3281	GGAACCCAAA	CTCACACCGC	TGTCGGAAAT	TTCTCAGAGA	CTCCTTAAAC	TGTACTCGGA	TAAATTTGGT	TCTGAAAATG 3360
3361	TCAAAATGAT	ACAGGATTCT	GGCAAGGTCA	ACCCTAAGGA	TCTGGATTCT	AAGTATGCAT	ACATCCAGGT	GACTCAAGTC 3440
3441	ATCCCCCTCT	TTGACGAAAA	AGAGTTGCAA	GAAAGGAAAA	CAGAGTTTGA	GAGATCCCAC	AACATCCGCC	GCTTCATGTT 3520
3521	TGAGATGCCA	TTTACGACGA	CCGGGAAGAG	GCAGGGCGGG	GTGGAAGAGC	AGTGCAAACG	GCGCACCATC	CTGACAGCCA 3600
3601	TACACTGCTT	CCCTTATGTG	AAGAAGCGCA	TCCCTGTCTAT	GTACCAGCAC	CACACTGACC	TGAACCCCAT	CGAGGTGGCC 3680
3681	ATTGACGAGA	TGAGTAAGAA	GGTGGCGGAG	CTCCGGCAGC	TGTGCTCCTC	GGCCGAGGTG	GACATGATCA	AACTGCAGCT 3760
3761	CAAACCTCCAG	GGCAGCGTGA	GTGTTCAGGT	CAATGCTGGC	CCACTAGCAT	ATGCGCGAGC	TTTCTTAGAT	GATACAAACA 3840
3841	CAAAGCGATA	TCCTGACAAAT	AAAGTGAAGC	TGCTTAAGGA	AGTTTTTCAGG	CAATTTGTGG	AAGCTTGCGG	TCAAGCCTTA 3920
3921	GCGGTAAACG	AACGTCTGAT	TAAAGAAGAC	CAGCTCGAGT	ATCAGGAAGA	AATGAAAGCC	AACTACAGGG	AAATGGCGAA 4000
4001	GGAGCTTTCT	GAAATCATGC	ATGAGCAGAT	CTGCCCCCTG	GAGGAGAAGA	CGAGCGTCTT	ACCGAATTCC	CTTCACATCT 4080
4081	TCAACGCCAT	CAGTGGGACT	CCAACAAGCA	CAATGGTTCA	CGGGATGACC	AGCTCGTCTT	CGGTCTGTGT	ATTACATCTC 4160
4161	ATGGCCCGTG	TGTGGGGACT	TGCTTTGTCA	TTTGCAAACT	CAGGATGCTT	TCCAAAGCCA	ATCACTGGGG	AGACCGAGCA 4240
4241	CAGGGAGGAC	CAAGGGGAAG	GGGAGAGAAA	GGAATAAAG	AACAACGTTA	TTTCTTAACA	GACTTTCTAT	AGGAGTTGTA 4320
4321	AGAAGGTGCA	CATATTTTTT	TAAATCTCAC	TGGCAATATT	CAAAGTTTTC	ATTGTGTCTT	AACAAAGGTG	TGGTAGACAC 4400
4401	TCTTGAGCTG	GACTTAGATT	TTATCTTCTC	TTGCAGAGTA	GTGTTAGAAT	AGATGGCCTA	CAGAAAAAAA	AGGTTCTGGG 4480
4481	ATCTACATGG	CAGGGAGGGC	TGCACTGACA	TTGATGCCTG	GGGGACCTTT	TGCTTCGACT	CGTGCCGGAA	ATCTGATCGT 4560
4561	AATCAGGGTA	CAGAACCTTAC	TAGTTTGTGC	TAGGAGTATG	TTGTATGACT	AGGATTTGTG	CTATTATCTC	ATTCAACAAC 4640
4641	ATAGAGCAAG	AATAGTGAGC	TAACTGAGCT	AGACACTCAA	TTAATCCGCT	ACTGGCTTCA	AGTCAGAACT	TTGTCAATTAA 4720
4721	TCATCGACTC	CGGGACGGTC	ATATATGTAT	TACATTTCTA	CATTTTAAAT	ACTCACATGG	GCTTATGCAT	TAAGTTTAAAT 4800
4801	TGTGATAAAT	TTGTGCTGGT	CCAGTATATG	CAATACACTT	TAATGGTTTA	TTCTTGTCTAT	AAAAATGTGC	AATATGGAGA 4880
4881	TGTATACAAG	TCTTTACT						4898

FIG. 10H

	10	20	30	40	50	60	70	80
1	MEGHVMI AFL	PTILNQLFRV	LTRATQEEVA	VNVTRVIIHV	VAQCHEEGLE	SHLRSYVKYA	YKAEPYVASE	YKTVHEELTK 80
81	SMTTILKPSA	DFLT SNKLLR	YSWFFFDVLI	KSMAQH LIEN	SKVKLLRNQR	FPASYHHA AE	TVVNMLMPHI	TQKFGDNPEA 160
161	SKNANHSLAV	FIKRCFTFMD	RGFVFKQINN	YISCFAPGDP	KTLFEYKFEP	LRVVCNHEHY	IPLNLPMPFG	KGRIQRYQDL 240
241	QLDYSLTDEF	CRNHFLVGLL	LREVG TALQE	FREVRLIAIS	VLKNLLIKHS	FDDRYASRSH	QARIATLYLP	LFGLLIENVQ 320
321	RINVRDVSPF	PVNAGMTVKD	ESLALPAVNP	LVTPOKGSTL	DNSLHKDLLG	AISG IASPYT	TSTPNINSVR	NADSRGSLIS 400
401	TDSGNSLPER	NSEKSNSLDK	HQQSSTLGNS	VVRCDKLDQS	EIKSLLMCFL	YILKSM SDDA	LFTYWNKAST	SELM DFFTIS 480
481	EVCLHQFOYM	GKRYIARNQE	GLGPIVHDRK	SQTLPVSRNR	TGMMHARLOQ	LGSLD NSLTF	NHSYGHSDAD	VLHQSLLEAN 560
561	IATEVCLTAL	DTLSLFTLAF	KNOLLADHGH	NPLMKKVFDV	YLCFLQKHQS	ETALKNVFTA	LRSLIYKFPS	TFYEGRADMC 640
641	AALCYEILKC	CNSKLSSIRT	EASQLLYFLM	RNNFDYTGKK	SFVRTHLQVI	ISVSQ LIADV	VGIGETRFOQ	SLSIINN CAN 720
721	SDRLIKHTSF	SSDVKDLTKR	IRTVLMATAQ	MKEHENDPEM	LVDLOYS LAK	SYASTPELRK	TWLD SMARIH	VKNGDLSEAA 800
801	MCYVHV TALV	AEYLTRKGVF	RQGCTAFRVI	TPNIDEEASM	MEDVGMQDVH	FNEDVLMELL	EQCADGLWKA	ERYELIADIY 880
881	KLIPIY EKR	RDFFEDEDGK	EYIYKEPKLT	PLSEISORLL	KLYSDKFGSE	NVTMIQDSGK	VNPKDLDSKY	AYIQVTHVIP 960
961	FFDEKELQER	KTEFERSHNI	RRFMFEMPFT	QTGKRQGGVE	EOCKRRTILT	AIHCFPYVKK	RIPVMYQHHT	DLNPIEVAID 1040
1041	EMSKKVAELR	QLCSSAEVDM	IKLQ LKQGS	VSVQVNAGPL	AYARAF LDDT	NTKRYPDNKV	KLLKEVFRQF	VEACQALAV 1120
1121	NERLIKEDQL	EYQEEMKANY	REMAKELSEI	MHEQICPLEE	KTSVLPNSLH	IFNAISGTPT	STMVHGMTSS	SSVV 1194

09687837-101300

FIG. 10H (cont.)

Exon 1A (-182 to -102)

GCAGGGGAAAAACCTGGCCCCATGATTCACCTACTTCCCACCGGATCTCTCCCATGACACGTGAGGATTA  
TTACAATTTAA -102

Exon 1B (-219 to -102)

TTATCCCTTTACTACTTGCGAAGTGAGTTCGGTAGATGGGAGTGGAGAAGAGAACCTTAGAATCATTGTTTAGTCTTCAT  
CTTTCACAGCTCAGGCTGAAGGCCTTTCCTTGCTGAGA -102

Exon 1C (-143 to -102)

GCGGCAGAGCGTGTCTGAGGTGGTGCGCGGCTCCGTGCTCCT -102

Exon2 and the rest of human CLASP2 cDNA

-101 -79  
GGCAAAGCCAAAGCTAATTGAGC

-78 -1  
AAGCTAATTGAGCCACTCGACTATGAAAATGTCATCGTCCAGAAGAAGACTCAGATCCTGAACGACTGTTTACGGGAG

1/1	31/11
ATG CTG CTC TTC CCT TAC GAT GAC TTT CAG	ACG GCC ATC CTG AGA CGA CAG GGT CGA TAC
Met leu leu phe pro tyr asp asp phe gln	thr ala ile leu arg arg gln gly arg tyr
61/21	91/31
ATA TGC TCA ACA GTG CCT GCG AAG GCG GAA	GAG GAA GCA CAG AGC TTG TTT GTT ACA GAG
ile cys ser thr val pro ala lys ala glu	glu glu ala gln ser leu phe val thr glu
121/41	151/51
TGC ATC AAA ACC TAT AAC TCT GAC TGG CAT	CTT GTG AAC TAT AAA TAT GAA GAT TAC TCA
cys ile lys thr tyr asn ser asp trp his	leu val asn tyr lys tyr glu asp tyr ser
181/61	211/71
GGA GAG TTT CGA CAG CTT CCG AAC AAA GTG	GTC AAG TTG GAT AAA CTT CCA GTT CAT GTC
gly glu phe arg gln leu pro asn lys val	val lys leu asp lys leu pro val his val
241/81	271/91
TAT GAA GTT GAC GAG GAG GTC GAC AAA GAT	GAG GAT GCT GCC TCC CTT GGT TCC CAG AAG
tyr glu val asp glu glu val asp lys asp	glu asp ala ala ser leu gly ser gln lys
301/101	331/111
GGT GGG ATC ACC AAG CAT GGC TGG CTG TAC	AAA GGC AAC ATG AAC AGT GCC ATC AGC GTG
gly gly ile thr lys his gly trp leu tyr	lys gly asn met asn ser ala ile ser val
361/121	391/131
ACC ATG AGG TCA TTT AAG AGA CGA TTT TTC	CAC CTG ATT CAA CTT GGC GAT GGA TCC TAT
thr met arg ser phe lys arg arg phe phe	his leu ile gln leu gly asp gly ser tyr
421/141	451/151
AAT TTG AAT TTT TAT AAA GAT GAA AAG ATC	TCC AAA GAA CCA AAA GGA TCA ATA TTT CTG
asn leu asn phe tyr lys asp glu lys ile	ser lys glu pro lys gly ser ile phe leu
481/161	511/171
GAT TCC TGT ATG GGT GTC GTT CAG AAC AAC	AAA GTC AGG CGT TTT GCT TTT GAG CTC AAG
asp ser cys met gly val val gln asn asn	lys val arg arg phe ala phe glu leu lys

541/181	571/191
ATG CAG GAC AAA AGT AGT TAT CTC TTG GCA	GCA GAC AGT GAA GTG GAA ATG GAA GAA TGG
met gln asp lys ser ser tyr leu leu ala	ala asp ser glu val glu met glu glu trp
601/201	631/211
ATC ACA ATT CTA AAT AAG ATC CTC CAG CTC	AAC TTT GAA GCT GCA ATG CAA GAA AAG CGA
ile thr ile leu asn lys ile leu gln leu	asn phe glu ala ala met gln glu lys arg
661/221	691/231
AAT GGC GAC TCT CAC GAA GAT GAT GAA CAA	AGC AAA TTG GAA GGT TCT GGT TCC GGT TTA
asn gly asp ser his glu asp asp glu gln	ser lys leu glu gly ser gly ser gly leu
721/241	751/251
GAT AGC TAC CTG CCG GAA CTT GCC AAG AGT	GCA AGA GAA GCA GAA ATC AAA CTA AAA AGT
asp ser tyr leu pro glu leu ala lys ser	ala arg glu ala glu ile lys leu lys ser
781/261	811/271
GAA AGC AGA GTC AAA CTT TTT TAT TTG GAC	CCA GAT GCC CAG AAG CTT GAC TTC TCA TCA
glu ser arg val lys leu phe tyr leu asp	pro asp ala gln lys leu asp phe ser ser
841/281	871/291
GCT GAG CCA GAA GTG AAG TCA TTT GAA GAG	AAG TTT GGA AAA AGG ATC CTT GTC AAG TGC
ala glu pro glu val lys ser phe glu glu	lys phe gly lys arg ile leu val lys cys
901/301	931/311
AAT GAT TTA TCT TTC AAT TTG CAA TGC TGT	GTT GCC GAA AAT GAA GAA GGA CCC ACT ACA
asn asp leu ser phe asn leu gln cys cys	val ala glu asn glu glu gly pro thr thr
961/321	991/331
AAT GTT GAA CCT TTC TTT GTT ACT CTA TCC	CTG TTT GAC ATA AAA TAC AAC CGG AAG ATT
asn val glu pro phe phe val thr leu ser	leu phe asp ile lys tyr asn arg lys ile
1021/341	1051/351
TCT GCC GAT TTC CAC GTA GAC CTG AAC CAT	TTC TCA GTG AGG CAA ATG CTC GCC ACC ACG
ser ala asp phe his val asp leu asn his	phe ser val arg gln met leu ala thr thr
1081/361	1111/371
TCC CCG GCG CTG ATG AAT GGC AGT GGG CAG	AGC CCA TCT GTC CTC AAG GGC ATC CTT CAT
ser pro ala leu met asn gly ser gly gln	ser pro ser val leu lys gly ile leu his
1141/381	1171/391
GAA GCC GCC ATG CAG TAT CCG AAG CAG GGA	ATA TTT TCA GTC ACT TGT CCT CAT CCA GAT
glu ala ala met gln tyr pro lys gln gly	ile phe ser val thr cys pro his pro asp
1201/401	1231/411
ATA TTT CTT GTG GCC AGA ATT GAA AAA GTC	CTT CAG GGG AGC ATC ACA CAT TGC GCT GAG
ile phe leu val ala arg ile glu lys val	leu gln gly ser ile thr his cys ala glu
1261/421	1291/431
CCA TAT ATG AAA AGT TCA GAC TCT TCT AAG	GTG GCC CAG AAG GTG CTG AAG AAT GCC AAG
pro tyr met lys ser ser asp ser ser lys	val ala gln lys val leu lys asn ala lys
1321/441	1351/451
CAG GCA TGC CAA AGA CTA GGA CAG TAT AGA	ATG CCA TTT GCT TGG GCA GCA AGG ACA TTG
gln ala cys gln arg leu gly gln tyr arg	met pro phe ala trp ala ala arg thr leu

1381/461  
TTT AAG GAT GCA TCT GGA AAT CTT GAC AAA AAT GCC AGA TTT TCT GCC ATC TAC AGG CAA  
phe lys asp ala ser gly asn leu asp lys asn ala arg phe ser ala ile tyr arg gln

1411/471  
1441/481  
GAC AGC AAT AAG CTA TCC AAT GAT GAC ATG CTC AAG TTA CTT GCA GAC TTT CGG AAA CCT  
asp ser asn lys leu ser asn asp asp met leu lys leu leu ala asp phe arg lys pro

1471/491  
1501/501  
GAG AAG ATG GCT AAG CTC CCA GTG ATT TTA GGC AAT CTA GAC ATT ACA ATT GAT AAT GTT  
glu lys met ala lys leu pro val ile leu gly asn leu asp ile thr ile asp asn val

1531/511  
1561/521  
TCC TCA GAC TTC CCT AAT TAT GTT AAT TCA TCA TAC ATT CCC ACA AAA CAA TTT GAA ACC  
ser ser asp phe pro asn tyr val asn ser ser tyr ile pro thr lys gln phe glu thr

1591/531  
1621/541  
TGC AGT AAA ACT CCC ATC ACG TTT GAA GTG GAG GAA TTT GTG CCC TGC ATA CCA AAA CAC  
cys ser lys thr pro ile thr phe glu val glu glu phe val pro cys ile pro lys his

1651/551  
1681/561  
ACT CAG CCT TAC ACC ATC TAC ACC AAT CAC CTT TAC GTT TAT CCT AAG TAC TTG AAA TAC  
thr gln pro tyr thr ile tyr thr asn his leu tyr val tyr pro lys tyr leu lys tyr

1711/571  
1741/581  
GAC AGT CAG AAG TCT TTT GCC AAG GCT AGA AAT ATT GCG ATT TGC ATT GAA TTC AAA GAT  
asp ser gln lys ser phe ala lys ala arg asn ile ala ile cys ile glu phe lys asp

1771/591  
1801/601  
TCA GAT GAG GAA GAC TCT CAG CCC CTT AAG TGC ATT TAT GGC AGA CCT GGT GGG CCA GTT  
ser asp glu glu asp ser gln pro leu lys cys ile tyr gly arg pro gly gly pro val

1831/611  
1861/621  
TTC ACA AGA AGC GCC TTT GCT GCA GTT TTA CAC CAT CAC CAA AAC CCA GAA TTT TAT GAT  
phe thr arg ser ala phe ala ala val leu his his his gln asn pro glu phe tyr asp

1891/631  
1921/641  
GAG ATT AAA ATA GAG TTG CCC ACT CAG CTG CAT GAA AAG CAC CAC CTG TTG CTC ACA TTC  
glu ile lys ile glu leu pro thr gln leu his glu lys his his leu leu leu thr phe

1951/651  
1981/661  
TTC CAT GTC AGC TGT GAC AAC TCA AGT AAA GGA AGC ACG AAG AAG AGG GAT GTC GTT GAA  
phe his val ser cys asp asn ser ser lys gly ser thr lys lys arg asp val val glu

2011/671  
2041/681  
ACC CAA GTT GGC TAC TCC TGG CTT CCC CTC CTG AAA GAC GGA AGG GTG GTG ACA AGC GAG  
thr gln val gly tyr ser trp leu pro leu leu lys asp gly arg val val thr ser glu

2071/691  
2101/701  
CAG CAC ATC CCG GTC TCG GCG TAC CTT CCT TCG GGC CAT CTT GGC TAC CAA GAG CTT GGG  
gln his ile pro val ser ala tyr leu pro ser gly his leu gly tyr gln glu leu gly

2131/711  
2161/721  
ATG GGC AGG CAT TAT GGT CCG GAA ATT AAA TGG GTA GAT GGA GGC AAG CCA CTG CTG AAA  
met gly arg his tyr gly pro glu ile lys trp val asp gly gly lys pro leu leu lys

2191/731

FIG. 11A  
3 of 8



2221/741 2251/751  
ATT TCC ACT CAT CTG GTT TCT ACA GTG TAT ACT CAG GAT CAG CAT TTA CAT AAT TTT TTC  
ile ser thr his leu val ser thr val tyr thr gln asp gln his leu his asn phe phe

2281/761 2311/771  
CAG TAC TGT CAG AAA ACC GAA TCT GGA GCC CAA GCC TTA GGA AAC GAA CTT GTA AAG TAC  
gln tyr cys gln lys thr glu ser gly ala gln ala leu gly asn glu leu val lys tyr

2341/781 2371/791  
CTT AAG AGT CTG CAT GCG ATG GAA GGC CAC GTG ATG ATC GCC TTC TTG CCC ACT ATC CTA  
leu lys ser leu his ala met glu gly his val met ile ala phe leu pro thr ile leu

2401/801 2431/811  
AAC CAG CTG TTC CGA GTC CTC ACC AGA GCC ACA CAG GAA GAA GTC GCG GTT AAC GTG ACT  
asn gln leu phe arg val leu thr arg ala thr gln glu glu val ala val asn val thr

2461/821 2491/831  
CGG GTC ATT ATT CAT GTG GTT GCC CAG TGC CAT GAG GAA GGA TTG GAG AGC CAC TTG AGG  
arg val ile ile his val val ala gln cys his glu glu gly leu glu ser his leu arg

2521/841 2551/851  
TCA TAT GTT AAG TAC GCG TAT AAG GCT GAG CCA TAT GTT GCC TCT GAA TAC AAG ACA GTG  
ser tyr val lys tyr ala tyr lys ala glu pro tyr val ala ser glu tyr lys thr val

2581/861 2611/871  
CAT GAA GAA CTG ACC AAA TCC ATG ACC ACG ATT CTC AAG CCT TCT GCC GAT TTC CTC ACC  
his glu glu leu thr lys ser met thr thr ile leu lys pro ser ala asp phe leu thr

2641/881 2671/891  
AGC AAC AAA CTA CTG AGG TAC TCA TGG TTT TTC TTT GAT GTA CTG ATC AAA TCT ATG GCT  
ser asn lys leu leu arg tyr ser trp phe phe phe asp val leu ile lys ser met ala

2701/901 2731/911  
CAG CAT TTG ATA GAG AAC TCC AAA GTT AAG TTG CTG CGA AAC CAG AGA TTT CCT GCA TCC  
gln his leu ile glu asn ser lys val lys leu leu arg asn gln arg phe pro ala ser

2761/921 2791/931  
TAT CAT CAT GCA GCG GAA ACC GTT GTA AAT ATG CTG ATG CCA CAC ATC ACT CAG AAG TTT  
tyr his his ala ala glu thr val val asn met leu met pro his ile thr gln lys phe

2821/941 2851/951  
GGA GAT AAT CCA GAG GCA TCT AAG AAC GCG AAT CAT AGC CTT GCT GTC TTC ATC AAG AGA  
gly asp asn pro glu ala ser lys asn ala asn his ser leu ala val phe ile lys arg

2881/961 2911/971  
TGT TTC ACC TTC ATG GAC AGG GGC TTT GTC TTC AAG CAG ATC AAC AAC TAC ATT AGC TGT  
cys phe thr phe met asp arg gly phe val phe lys gln ile asn asn tyr ile ser cys

2941/981 2971/991  
TTT GCT CCT GGA GAC CCA AAG ACC CTC TTT GAA TAC AAG TTT GAA TTT CTC CGT GTA GTG  
phe ala pro gly asp pro lys thr leu phe glu tyr lys phe glu phe leu arg val val

3001/1001 3031/1011  
TGC AAC CAT GAA CAT TAT ATT CCG TTG AAC TTA CCA ATG CCA TTT GGA AAA GGC AGG ATT  
cys asn his glu his tyr ile pro leu asn leu pro met pro phe gly lys gly arg ile



3061/1021	3091/1031
CAA AGA TAC CAA GAC CTC CAG CTT GAC TAC TCA TTA ACA GAT GAG TTC TGC AGA AAC CAC	
gln arg tyr gln asp leu gln leu asp tyr ser leu thr asp glu phe cys arg asn his	
3121/1041	3151/1051
TTC TTG GTG GGA CTG TTA CTG AGG GAG GTG GGG ACA GCC CTC CAG GAG TTC CGG GAG GTC	
phe leu val gly leu leu leu arg glu val gly thr ala leu gln glu phe arg glu val	
3181/1061	3211/1071
CGT CTG ATC GCC ATC AGT GTG CTC AAG AAC CTG CTG ATA AAG CAT TCT TTT GAT GAC AGA	
arg leu ile ala ile ser val leu lys asn leu leu ile lys his ser phe asp asp arg	
3241/1081	3271/1091
TAT GCT TCA AGG AGC CAT CAG GCA AGG ATA GCC ACC CTC TAC CTG CCT CTG TTT GGT CTG	
tyr ala ser arg ser his gln ala arg ile ala thr leu tyr leu pro leu phe gly leu	
3301/1101	3331/1111
CTG ATT GAA AAC GTC CAG CGG ATC AAT GTG AGG GAT GTG TCA CCC TTC CCT GTG AAC GCG	
leu ile glu asn val gln arg ile asn val arg asp val ser pro phe pro val asn ala	
3361/1121	3391/1131
GGC ATG ACC GTG AAG GAT GAA TCC CTG GCT CTA CCA GCT GTG AAT CCG CTG GTG ACG CCG	
gly met thr val lys asp glu ser leu ala leu pro ala val asn pro leu val thr pro	
3421/1141	3451/1151
CAG AAG GGA AGC ACC CTG GAC AAC AGC CTG CAC AAG GAC CTG CTG GGC GCC ATC TCC GGC	
gln lys gly ser thr leu asp asn ser leu his lys asp leu leu gly ala ile ser gly	
3481/1161	3511/1171
ATT GCT TCT CCA TAT ACA ACC TCA ACT CCA AAC ATC AAC AGT GTG AGA AAT GCT GAT TCG	
ile ala ser pro tyr thr thr ser thr pro asn ile asn ser val arg asn ala asp ser	
3541/1181	3571/1191
AGA GGA TCT CTC ATA AGC ACA GAT TCG GGT AAC AGC CTT CCA GAA AGG AAT AGT GAG AAG	
arg gly ser leu ile ser thr asp ser gly asn ser leu pro glu arg asn ser glu lys	
3601/1201	3631/1211
AGC AAT TCC CTG GAT AAG CAC CAA CAA AGT AGC ACA TTG GGA AAT TCC GTG GTT CGC TGT	
ser asn ser leu asp lys his gln gln ser ser thr leu gly asn ser val val arg cys	
3661/1221	3691/1231
GAT AAA CTT GAC CAG TCT GAG ATT AAG AGC CTA CTG ATG TGT TTC CTC TAC ATC TTA AAG	
asp lys leu asp gln ser glu ile lys ser leu leu met cys phe leu tyr ile leu lys	
3721/1241	3751/1251
AGC ATG TCT GAT GAT GCT TTG TTT ACA TAT TGG AAC AAG GCT TCA ACA TCT GAA CTT ATG	
ser met ser asp asp ala leu phe thr tyr trp asn lys ala ser thr ser glu leu met	
3781/1261	3811/1271
GAT TTT TTT ACA ATA TCT GAA GTC TGC CTG CAC CAG TTC CAG TAC ATG GGG AAG CGA TAC	
asp phe phe thr ile ser glu val cys leu his gln phe gln tyr met gly lys arg tyr	
3841/1281	3871/1291
ATA GCC AGG AAC CAG GAG GGG TTG GGA CCC ATA GTT CAT GAT CGA AAG TCT CAG ACA TTG	
ile ala arg asn gln glu gly leu gly pro ile val his asp arg lys ser gln thr leu	

3901/1301 3931/1311  
CCT GTT TCC CGT AAC AGA ACA GGA ATG ATG CAT GCC AGA TTG CAG CAG CTG GGC AGC CTG  
pro val ser arg asn arg thr gly met met his ala arg leu gln gln leu gly ser leu

3961/1321 3991/1331  
GAT AAC TCT CTC ACT TTT AAC CAC AGC TAT GGC CAC TCG GAC GCA GAT GTT CTG CAC CAG  
asp asn ser leu thr phe asn his ser tyr gly his ser asp ala asp val leu his gln

4021/1341 4051/1351  
TCA TTA CTT GAA GCC AAC ATT GCT ACT GAG GTT TGC CTG ACA GCT CTG GAC ACG CTT TCT  
ser leu leu glu ala asn ile ala thr glu val cys leu thr ala leu asp thr leu ser

4081/1361 4111/1371  
CTA TTT ACA TTG GCG TTT AAG AAC CAG CTC CTG GCC GAC CAT GGA CAT AAT CCT CTC ATG  
leu phe thr leu ala phe lys asn gln leu leu ala asp his gly his asn pro leu met

4141/1381 4171/1391  
AAA AAA GTT TTT GAT GTC TAC CTG TGT TTT CTT CAA AAA CAT CAG TCT GAA ACG GCT TTA  
lys lys val phe asp val tyr leu cys phe leu gln lys his gln ser glu thr ala leu

4201/1401 4231/1411  
AAA AAT GTC TTC ACT GCC TTA AGG TCC TTA ATT TAT AAG TTT CCC TCA ACA TTC TAT GAA  
lys asn val phe thr ala leu arg ser leu ile tyr lys phe pro ser thr phe tyr glu

4261/1421 4291/1431  
GGG AGA GCG GAC ATG TGT GCG GCT CTG TGT TAC GAG ATT CTC AAG TGC TGT AAC TCC AAG  
gly arg ala asp met cys ala ala leu cys tyr glu ile leu lys cys cys asn ser lys

4321/1441 4351/1451  
CTG AGC TCC ATC AGG ACG GAG GCC TCC CAG CTG CTC TAC TTC CTG ATG AGG AAC AAC TTT  
leu ser ser ile arg thr glu ala ser gln leu leu tyr phe leu met arg asn asn phe

4381/1461 4411/1471  
GAT TAC ACT GGA AAG AAG TCC TTT GTC CGG ACA CAT TTG CAA GTC ATC ATA TCT GTC AGC  
asp tyr thr gly lys lys ser phe val arg thr his leu gln val ile ile ser val ser

4441/1481 4471/1491  
CAG CTG ATA GCA GAC GTT GTT GGC ATT GGG GAA ACC AGA TTC CAG CAG TCC CTG TCC ATC  
gln leu ile ala asp val val gly ile gly glu thr arg phe gln gln ser leu ser ile

4501/1501 4531/1511  
ATC AAC AAC TGT GCC AAC AGT GAC CGG CTT ATT AAG CAC ACC AGC TTC TCC TCT GAT GTG  
ile asn asn cys ala asn ser asp arg leu ile lys his thr ser phe ser ser asp val

4561/1521 4591/1531  
AAG GAC TTA ACC AAA AGG ATA CGC ACG GTG CTA ATG GCC ACC GCC CAG ATG AAG GAG CAT  
lys asp leu thr lys arg ile arg thr val leu met ala thr ala gln met lys glu his

4621/1541 4651/1551  
GAG AAC GAC CCA GAG ATG CTG GTG GAC CTC CAG TAC AGC CTG GCC AAA TCC TAT GCC AGC  
glu asn asp pro glu met leu val asp leu gln tyr ser leu ala lys ser tyr ala ser

4681/1561 4711/1571  
ACG CCC GAG CTC AGG AAG ACG TGG CTC GAC AGC ATG GCC AGG ATC CAT GTC AAA AAT GGC  
thr pro glu leu arg lys thr trp leu asp ser met ala arg ile his val lys asn gly

4741/1581	4771/1591
GAT CTC TCA GAG GCA GCA ATG TGC TAT GTC	CAC GTA ACA GCC CTA GTG GCA GAA TAT CTC
asp leu ser glu ala ala met cys tyr val	his val thr ala leu val ala glu tyr leu
4801/1601	4831/1611
ACA CGG AAA GGC GTG TTT AGA CAA GGA TGC	ACC GCC TTC AGG GTC ATT ACC CCA AAC ATC
thr arg lys gly val phe arg gln gly cys	thr ala phe arg val ile thr pro asn ile
4861/1621	4891/1631
GAC GAG GAG GCC TCC ATG ATG GAA GAC GTG	GGG ATG CAG GAT GTC CAT TTC AAC GAG GAT
asp glu glu ala ser met met glu asp val	gly met gln asp val his phe asn glu asp
4921/1641	4951/1651
GTG CTG ATG GAG CTC CTT GAG CAG TGC GCA	GAT GGA CTC TGG AAA GCC GAG CGC TAC GAG
val leu met glu leu leu glu gln cys ala	asp gly leu trp lys ala glu arg tyr glu
4981/1661	5011/1671
CTC ATC GCC GAC ATC TAC AAA CTT ATC ATC	CCC ATT TAT GAG AAG CGG AGG GAT TTC TTT
leu ile ala asp ile tyr lys leu ile ile	pro ile tyr glu lys arg arg asp phe phe
5041/1681	5071/1691
GAA GAT GAA GAT GGA AAG GAG TAT ATT TAC	AAG GAA CCC AAA CTC ACA CCG CTG TCG GAA
glu asp glu asp gly lys glu tyr ile tyr	lys glu pro lys leu thr pro leu ser glu
5101/1701	5131/1711
ATT TCT CAG AGA CTC CTT AAA CTG TAC TCG	GAT AAA TTT GGT TCT GAA AAT GTC AAA ATG
ile ser gln arg leu leu lys leu tyr ser	asp lys phe gly ser glu asn val lys met
5161/1721	5191/1731
ATA CAG GAT TCT GGC AAG GTC AAC CCT AAG	GAT CTG GAT TCT AAG TAT GCA TAC ATC CAG
ile gln asp ser gly lys val asn pro lys	asp leu asp ser lys tyr ala tyr ile gln
5221/1741	5251/1751
GTG ACT CAC GTC ATC CCC TTC TTT GAC GAA	AAA GAG TTG CAA GAA AGG AAA ACA GAG TTT
val thr his val ile pro phe phe asp glu	lys glu leu gln glu arg lys thr glu phe
5281/1761	5311/1771
GAG AGA TCC CAC AAC ATC CGC CGC TTC ATG	TTT GAG ATG CCA TTT ACG CAG ACC GGG AAG
glu arg ser his asn ile arg arg phe met	phe glu met pro phe thr gln thr gly lys
5341/1781	5371/1791
AGG CAG GGC GGG GTG GAA GAG CAG TGC AAA	CGG CGC ACC ATC CTG ACA GCC ATA CAC TGC
arg gln gly gly val glu glu gln cys lys	arg arg thr ile leu thr ala ile his cys
5401/1801	5431/1811
TTC CCT TAT GTG AAG AAG CGC ATC CCT GTC	ATG TAC CAG CAC CAC ACT GAC CTG AAC CCC
phe pro tyr val lys lys arg ile pro val	met tyr gln his his thr asp leu asn pro
5461/1821	5491/1831
ATC GAG GTG GCC ATT GAC GAG ATG AGT AAG	AAG GTG GCG GAG CTC CGG CAG CTG TGC TCC
ile glu val ala ile asp glu met ser lys	lys val ala glu leu arg gln leu cys ser
5521/1841	5551/1851
TCG GCC GAG GTG GAC ATG ATC AAA CTG CAG	CTC AAA CTC CAG GGC AGC GTG AGT GTT CAG
ser ala glu val asp met ile lys leu gln	leu lys leu gln gly ser val ser val gln

5581/1861	5611/1871
GTC AAT GCT GGC CCA CTA GCA TAT GCG CGA	GCT TTC TTA GAT GAT ACA AAC ACA AAG CGA
val asn ala gly pro leu ala tyr ala arg	ala phe leu asp asp thr asn thr lys arg
5641/1881	5671/1891
TAT CCT GAC AAT AAA GTG AAG CTG CTT AAG	GAA GTT TTC AGG CAA TTT GTG GAA GCT TGC
tyr pro asp asn lys val lys leu leu lys	glu val phe arg gln phe val glu ala cys
5701/1901	5731/1911
GGT CAA GCC TTA GCG GTA AAC GAA CGT CTG	ATT AAA GAA GAC CAG CTC GAG TAT CAG GAA
gly gln ala leu ala val asn glu arg leu	ile lys glu asp gln leu glu tyr gln glu
5761/1921	5791/1931
GAA ATG AAA GCC AAC TAC AGG GAA ATG GCG	AAG GAG CTT TCT GAA ATC ATG CAT GAG CAG
glu met lys ala asn tyr arg glu met ala	lys glu leu ser glu ile met his glu gln
5821/1941	5851/1951
ATC TGC CCC CTG GAG GAG AAG ACG AGC GTC	TTA CCG AAT TCC CTT CAC ATC TTC AAC GCC
ile cys pro leu glu glu lys thr ser val	leu pro asn ser leu his ile phe asn ala
5881/1961	5911/1971
ATC AGT GGG ACT CCA ACA AGC ACA ATG GTT	CAC GGG ATG ACC AGC TCG TCT TCG GTC GTG
ile ser gly thr pro thr ser thr met val	his gly met thr ser ser ser ser val val
5941/1981	5971
TGA TTA CAT CTC ATG GCC CGT GTG TGG GGA	CTT GCT TTG TCA TTT GCA AAC TCA GGA TGC
5971	
6001	6031
TTT CCA AAG CCA ATC ACT GGG GAG ACC GAG	CAC AGG GAG GAC CAA GGG GAA GGG GAG AGA
6061	6091
AAG GAA ATA AAG AAC AAC GTT ATT TCT TAA	CAG ACT TTC TAT AGG AGT TGT AAG AAG GTG
6121	6151
CAC ATA TTT TTT TAA ATC TCA CTG GCA ATA	TTC AAA GTT TTC ATT GTG TCT TAA CAA AGG
6181	6211
TGT GGT AGA CAC TCT TGA GCT GGA CTT AGA	TTT TAT TCT TCC TTG CAG AGT AGT GTT AGA
6241	6271
ATA GAT GGC CTA CAG AAA AAA AAG GTT CTG	GGA TCT ACA TGG CAG GGA GGG CTG CAC TGA
6301	6331
CAT TGA TGC CTG GGG GAC CTT TTG CCT CGA	CTC GTG CCG GAA ATC TGA TCG TAA TCA GGG
6361	6391
TAC AGA ACT TAC TAG TTT TGT CTA GGA GTA	TGT TGT ATG ACT AGG ATT TGT GCT ATT ATC
6421	6451
TCA TTC AAC AAC ATA GAG CAA GAA TAG TGA	GCT AAC TGA GCT AGA CAC TCA ATT AAT CCG
6481	6511
CTA CTG GCT TCA AGT CAG AAC TTT GTC ATT	AAT CAT CGA CTC CGG GAC GGT CAT ATA TGT
6541	6571
ATT ACA TTT CTA CAT TTT TAA TAC TCA CAT	GGG CTT ATG CAT TAA GTT TAA TTG TGA TAA
6601	6631
ATT TGT GCT GGT CCA GTA TAT GCA ATA CAC	TTT AAT GGT TTA TTC TTG TCA TAA AAA TGT
6661	
GCA ATA TGG AGA TGT ATA CAA GTC TTT ACT	

09687837-101300

A. Allelic variations: single nucleotide changes (polymorphism) between CLASP-2 cDNA isoforms

Isoform	Difference	Nucleotide(s)	Consequence
1	polymorphism	862	A to G change; mis-sense mutation
2	polymorphism		A to C change; mis-sense mutation changing codon from histidine to proline
3	polymorphism	2210	A to G change; mis-sense mutation changing codon from asparagine to glutamic acid
4	polymorphism	2225	C to T change; mis-sense mutation changing codon from histidine to tyrosine

B. Alternative splices

Isoform	Difference	Nucleotide(s)	Consequence
1	exon deletion	209-291	premature, in-frame stop codon leading to the production of a truncated, most likely soluble protein

These differences may be found separately or together in various combinations in the different human CLASP-2 isoforms

FIG. 11B

human CLASP-2

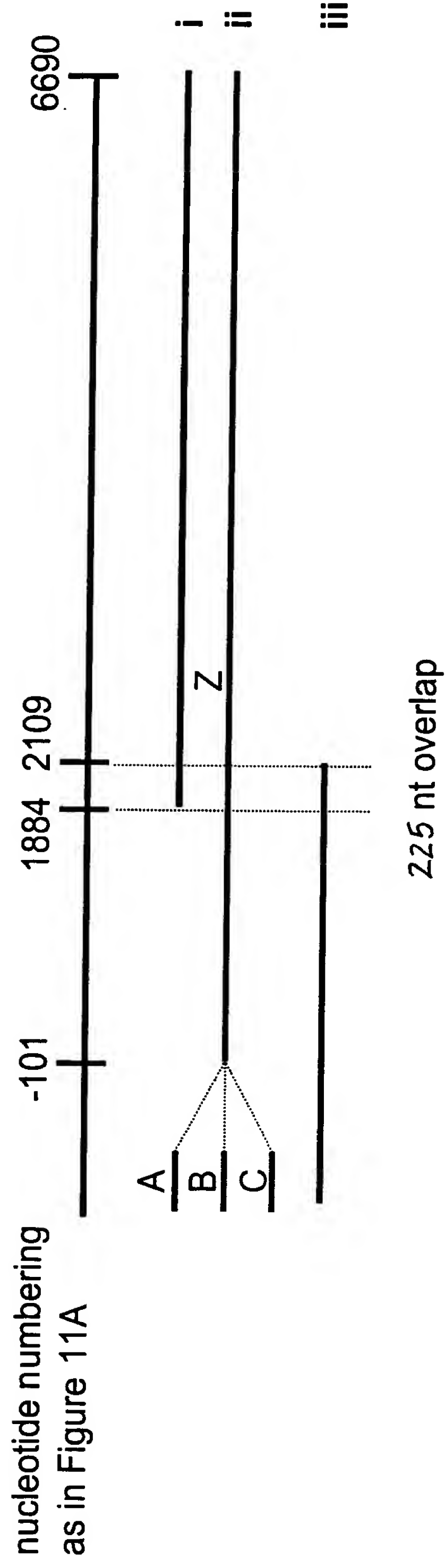


FIG. 11C



1st exon (nucleotides 335 to 445)

TGTCTTGCTTATCTTTTCGCCCTCCAGGCAAAGCCAAAGCTAATTGAGCCACT  
CGACTATGAAAATGTCATCGTCCAGAAGAAGACTCAGATCCTGAACGACTGT  
TTACGGGAGATGCTGCTCTTCCCTTACGATGACTTTCAGGTAAGTAACGTTAT  
GTTTCTATCCGTAGAACCACG

2nd exon (nucleotides 7101-7190)

TTACCCAAGGCTTTTCCTCCTGTTTTTGTTCAGACGGCCATCCTGAGACGA  
CAGGGTCGATACATATGCTCAACAGTGCCTGCGAAGGCGGAAGAGGAAGCA  
CAGAGCTTGTTTGTTACAGAGGTAAGGCTCTTTCCTGCATTAATTTACATTTT  
GAAGTCATTTTCCCCTAACTGCCTCC

3rd exon (nucleotides 11439 to 11521)

TTTTCTATTTTAAAAATCCCCCTTCAATAGTGTCATCAAAACCTATAACTCTGAC  
TGGCATCTTGTTGAACATAAATATGAAGATTACTCAGGAGAGTTTCGACAGC  
TTCCGAAGTGAGTAAGCTATATTATACACATAGGGGAAAAGTCTTT

4th exon (nucleotides 13987 to 14056)

CTAAACAAATTTTCTTTGTTGTTTTTATAGCAAAGTGGTCAAGTTGGATAAA  
CTTCCAGTTCATGTCTATGAAGTTGACGAGGAGGTGACAAAGATGAGGTGG  
GATACCTGCTTGCTGTTGCTTCTCTTTTCACTCTAGATTAA

5th exon (nucleotides 15212 to 15307)

GGAGGTTGACTGCTGGTGTTCCTTCTCTCCTAGGATGCTGCCTCCCTTGGTT  
CCCAGAAGGGTGGGATCACCAAGCATGGCTGGCTGTACAAAGGCAACATGA  
ACAGTGCCATCAGCGTGACCATGAGGGTGAGGACGCACATCACTTTGCCCTC  
CCCTCTCACAAGCCCTTTC

6th exon (nucleotides 16269 to 16404)

TGAAAGAATAGCTGTGTGTATATTTTTCTCTCAGTCATTTAAGAGACGATTTT  
TCCACCTGATTCAACTTGGCGATGGATCCTATAATTTGAATTTTATAAAGAT  
GAAAAGATCTCCAAAGAACCAAAAAGGATCAATATTTCTGGATTCCTGTATGG  
GTGTCGTTCAAGGTAAATATGAAAAGAGTTTTACCATTATGTTTTCTTA

7th exon (nucleotides 19459 to 19633)

AAGTATGTCTGTTTATCCTTTTTTTCATTTTCAGAACAAACAAAGTCAGGCGTTTT  
GCTTTTGAGCTCAAGATGCAGGACAAAAGTAGTTATCTCTTGGCAGCAGACA  
GTGAAGTGGAATGGAAGAATGGATCACAATTCTAAATAAGATCCTCCAGCT  
CAACTTTGAAGCTGCAATGCAAGAAAAGCGAAATGGCGACTCTCACGAAGGT  
AGATAGGCTTGGCTTCCCCCAGGCACATACACTCT

8th exon (nucleotides 20567 to 20634)

ATTACAAGTGATTCCGATAATCTGTTTTGCCATTTTAGATGATGAACAAAGCA  
AATTGGAAGGTTCTGGTTCCGGTTTAGATAGCTACCTGCCGGAAGTTGCCAAG  
GTAACATCGTCTTATATCTTCTGCTCTTCGTTGAATGC

9th exon (nucleotides 30257 to 30331)

GATTGTGTTAAATGTAATTTTCATGTATCTTGTTATCAGAGTGCAAGAGAAGC  
AGAAATCAAACATAAAAAGTGAAAGCAGAGTCAAACCTTTTTTATTTGGACCCA  
GATGCCCAGGTAAGAACTATCTAAATGTTTAATATTTAAAACCAAAT

10th exon (nucleotides 31851 to 31991)

CATAACTTATTTATATGTTTACATTTTCTTTTAAAGAAGCTTGACTTCTCATCA  
GCTGAGCCAGAAGTGAAGTCATTTGAAGAGAAGTTTGGAAAAAGGATCCTTG  
TCAAGTGCAATGATTTATCTTTCAATTTGCAATGCTGTGTTGCCGAAAATGAA  
GAAGGACCCACTACAAATGTAATTTTTCATTTTAAAAATAAACATTAAAAAA  
AAAATAGGCAG

11th exon (nucleotides 32472 to 32675)

CCATGGTGATCATTGGATTGTTTTGTTTTGTTTCAGGTTGAACCTTTCTTTGTTA  
CTCTATCCCTGTTTGACATAAAATACAACCGGAAGATTTCTGCCGATTTCCAC  
GTAGACCTGAACCATTTCTCAGTGAGGCAAATGCTCGCCACCACGTCCCCGG  
CGCTGATGAATGGCAGTGGGCAGAGCCCATCTGTCCTCAAGGGCATCCTTCA  
TGAAGCCGCCATGCAGTATCCGAAGCAGGTGGGGAGTATGAGCCCAGCATT  
CCTACTCAGACTCACTTTGCATGC

12th exon (nucleotides 33063 to 33185)

GAATTCTGCTTACTGAAGAAAATTGTTTGCCTCCTAGGGAATATTTTCAGTCA  
CTTGTCCTCATCCAGATATATTTCTTGTGGCCAGAATTGAAAAAGTCCTTCAG  
GGGAGCATCACACATTGCGCTGAGCCATATATGAAAAGTTTCAGACTCTTCTA  
AGGTATGAATGGCTTTTACGCTTTGGGGTGGTAAAAAGCAATCTGAA

13th exon (nucleotides 36702 to 36784)

CAGTATCTCATAGCTTTATTCTCATGTCTTCAAGGTGGCCCAGAAGGTGCTGA  
AGAATGCCAAGCAGGCATGCCAAAGACTAGGACAGTATAGAATGCCATTTGC  
TTGGGCAGCAAGGTAAGGAACACCTTTTATACCTTTTAAATCGATATAGATA  
GGTGATGG

14th partial exon (nucleotides 37353 to 37475)

GAAACCCAGTTTAGAAATGTTGCTTTGCCATTTTCAGGACATTGTTTAAGGATG  
CATCTGGAAATCTTGACAAAAATGCCAGATTTTCTGCCATCTACAGGCAAGA  
CAGCAATAAGCTATCCAATGATGACATGCTCAAGTTACTTGCAGACTTTTCGG  
AA

1 TACCAAGGGCAACTCTGGCACACCCTAAAGTCTGGAAAGGGGACATAGCTAGTCAGGGATGACCCGAGAAATGACTGGAAGCTCCACCAGAA  
93 TGCAGAGCTTCCTTTGTGCTTAAATAACTGAACAAGCATCACTCTGTGTAGCAGGACACCACCCAGCATTTTTTTGTCCCTTTGGAAACA  
185 CTTATTTCTGTTTCTTTGTGATACCAAACTAGCATACTCTAATTGTAGAAAATACAAAACATAGAGTAGAACATACTAAGTTCTTTATCTT  
277 AAGAAATGGCATTGTGTATGAGAATGTCTTGCTTATCTTTTCGCCCTCCAGGCAAAGCCAAAGCTAATTGAGCCACTCGACTATGAAAATG  
369 TCATCGTCCAGAAGAAGACTCAGATCCTGAACGACTGTTTACGGGAGATGCTGCTCTCCCTTACGATGACTTTCAGGTAAGTAACGTTATG  
461 TTTCTATCCGTAGAACACGTTGATCTTAACAAGCAGTATTTTTCTATGTATTGATTTATTGTTTGGTTAGTTAATTATTATTATTATT  
553 ATTTTTTTTGAGACACAGTCTTGCTCTGTGCACGCAGGCTTGAGTGCAGTGGTGCCATCTTGGCTCAACGGCAACCTCCGCCCTCTGGGTTCA  
645 AGCATTTCTCCTGCCTCAGCCTCTCAAGTAACTGGGATTACAGGCGTGTGCCACCATGCCTGGCTAATTTTTGTCTTTGTATTAGAGACAGG  
737 GTTTTGGCCACGGTGGCCAGGGTCGTCCTCAAACTCCTGGCCTTAAGTGATCTATCTGCCTTGGCCTCTCAAATGTTGGGATTATAGGCATGA  
829 GCCACTGTGCCCCGCCCTAATTATGGTTTTTAAAAGATGAAAATAAGATGTTATTTAAGAAAGAAAAGTTATTTTATATTCTTCCAAGCATCC  
921 TTCATGAGTTGATAATTTTTAATGGTATTATTTTTGCATATTAATTATAAGTATGCCAAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT  
1013 GTGTATCTTGAATAAAAGTGCTATACTCTGTCTGGGCTGATTTAGTGGGCACAAGTGCCCTTTCTGTCTTTTGAGATTTGTTTGTATTAGAT  
1105 TTTTTGGCCAAGTACCTTAGAATCTTAATGATGGGCTTTGCTGGGGAGCAGTGGGAGATTTCGTATGCCCTTTTTTTAGGATGGCATTGGG  
1197 AGCCTGCCTTCAGAGGCCCTGCACTCTGTGTGGGCTCCAGCAAAGCGTTCAAGGTTAGCCAAGAATGGCCTGAAGTTTACCTCTGTAGTGT  
1289 AATGTGGGTGCTGTTCCCTTGAAGAAATGTGGAGGACTCAGCACAGCTCCTGCTGTGTGCCGCTCTTTCAGGCTATGGCCTGTGGTTAAGAG  
1381 ACTAACAAGAAGCTGTGAGGCTGTTGAGGAATGAGAATGACATTCTTCCCTCCAGGAAACCCGGTGGTGTAAATGCCTTGCAGCGAGCCACC  
1473 TTGGTCCATTTGGAGGTTCTGGTTACTTTCCCTTGCTCTCCTGGGACCTGATCCTGGCACTTCTTTCTCCTCTCCTTTGATGCTCTTAGTTGG  
1565 ACACCTCTCTACCTGATGCTGTTACCATTAAAGCCCTGCTTTTTGTGAATCGAGCTGCCTTTTTTTTTTAAGCTTCACTGATTTCTTTGTTGT  
1657 TTGATTCCAAAAGTGTTACATCCATGTACAAAAAGATAAATGAGAGGGAAATATTGAAATAATTGACATGAAAAGCCTCCCCACGCCTTCTA  
1749 ATCCCATCCCACAGAACATGACTTAACTGTATACAGCTCTGTGCATACCTTGTCTTTAGAAACTTCCATGTTAATAGAAATTGTTAAATTACG  
1841 ATCCTTGAAGTTTTTTCCCAACCAATTTAAGCGACTCCAGCTTACAACAGAGGTGAGAATTTTACAAAATGTTCACTCTTTCTAACTTGT  
1933 AGAGATACCTGGGCCCCAAAATGATTATTCTTTAGCTCTGTCTGCATAAAAGGAATGCCCATGGGAATGAAATTGACCATTCTGTGTGGTGT  
2025 TGCTACCAAAGTAACAGGTAAATGGGTTGAGGTCATGCCAAACAATACCATGCTTTGCATACCTTCACTTTCATGACTAACTGCATGGGAACG  
2117 GACTAATAAATGAGAACCCTCTGAATGATGCCTTTTGCTGTGATTTGGCAACAAATGAAAAGCAAATCAAATGATTATAAATTGTACTGCA  
2209 TGTTGACAAGATTTTCTGTAGTGTGTCTGAGGAAGCTAAAGGTTATCTCAAATTTCTCTCAACATGAAGTATGTGTTCTTCTTGGTATTA  
2301 ATTAAGTAACAACCTTTTTTGAGTTTGAACCTAGAATGAAAAATTCTATTTGTATGACTGAGATAAAATTGCTTAAGAAACAACCAAGAA  
2393 ACGAGATACAGTTAGTTGAGTGTCTCTTTATCCAGGGAACAGGTATCTGGATGTTTAAGCAGTTGCAGAATCAGACAGTTTAACTTTGA  
2485 GAAAACCTCTGTGTCCCTTTGCTTTTAACTACTCTGGTGATAGCAGGCACAAATATTCTAGGAAAGGCAAAGAACTCACTAGCATTTTGT  
2577 GGCTAAGGTGATGAGCAAATATTATTTCTGTTTGGGGAGAAGTTTTCTAGAGATTTAGGAGCTTGAATTGGAGCTTTAATCCTCATCACA  
2669 GGAATTGTGATGGGCCCCAGTGAAGTTGGGTACAATTATTTGTTTTCTTATAGACTCCCACTTTCTTATCAGGTAAAGCCATGTACTCTGT  
2761 GCTTTCTTGTAAATTGTCTCAGTGATGTTATTAAGTGTCTAATTAGCTGGATGAGTGAAAGGCTTAAACAGTGCCACAGATTCTTTCTATC  
2853 TGTGTTTTCTTAGGCAGAATAAGAGCAGAATTATTGTATTATTAGAGGCAGAGGGAACAAATTAGATTGGGGAAAGTGTTTTATTTTCATATG  
2945 GAAAAGTAATACCAAGTTGGTTAGGAAATGGCAGCAGCAAACGCATGCTGAGGGGTGATTTACTGCCTTAAATAATTTAGCAGTATAAGT  
3037 TAACTATTAAAATAATAGAATTGGTGTCCATTTCTGCCAAATATATTTGAAATGACAATTTACTAAAATATAAGCATGGATAGTGGTGATG  
3129 CTTGTGTACATTTTCAAGTAGGCACATGTTGATCTTGAGCCTTTACTGGTCAGATCCTAAAGGCATCTACATGTTCTCTAAAATGAGTTG  
3221 TGTCAAGAAAAGATTTGCGGGTTGCATGTAGTTGCCCTGAGGATGACAGAAGAGTAGTTACTACAACAGCAGCAAAGAAGAGAGACATGAAGT  
3313 AAACGTGGATTTTTTAAAATCAAAGAATAGGCCAGGCGCACTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCAGATC  
3405 ACAAGGTCAGGAGTTTGAGACCAGCCTGGCCAATATGGTGAAACCCCATCTCTACCAGAAAATACAAAATTAGCCAGGCATGGTGGTGCATG  
3497 CCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCTCTTGAAACCAGGAGACAGAGGTTGCAGTAAGCTGAGATCGTGCCACTGCA  
3589 TTCCAGCCTGGGCGACAGAGTGAGACTCCATCTCAAAACAAACAAAAATCAAATAATAGTTCCCAGCCATCAGGTTATTGATGAAGTAGG  
3681 CTGGGCACGGTGGCTCACACCTGTAATCCCAGCACATTGGGAGTCCGAGGCAGGTGGATCACCTGAGGTGAGGTGTTTGGAGCCAGCCTGGC  
3773 CAACATGGCAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCAGGCATGGTGGTGGGCACCTATAATCCCAGCTACTTGGGAGGCTGAG  
3865 GCAAGAGAATCGCTTGAACCTGGGAGGTGGAGGTGCACTGAGCCAAGATCGCGCCATTGCACTCCAGCCTGGGGGACAAGAGGAAAACCTCC  
3957 ATCTCAAAAAAAAAGGGAATATTAATGAAGTAAAGTACATGTGATCTGCCATGGCCAGGGACAGGAATGCCATGGGGCCTGCAGCCGTCA  
4049 CTAGCTGATGGCCCTTCTTTTTGCAGAATCAGATCCTGTGCTTGGGGATCTCTGCCATCTGTGCTTTGGCTTCATGGTTCTCCTTGCCAGC  
4141 AGCATCTTCTCTTAGATCTTTCCCTACCTTTAGAGACCCTTGAAATCCCATATTGTCTGAAGCTATTTAAGTCCACAGAACTTTTCCC  
4233 CCCACTGTCTCAATTCCCTTTCTACTGCCTGTCTGCACCGTGCACATAAACACTTGAGTATGTGGTCTTGGCTGTTTACGACCTACTTCTTA  
4325 GGCTTCTTGCACGCAGGCATCCCGCCCCGTGCTGTGGTCTGAGAAGGGCTGGCTTTGAGCCTCTGTTCTCCACCCACCTGCCACCTACA  
4417 CATGCACAAAATCCCTTTCTTGCTAGGTGCTAGGGTTGAATACCCATTGCTTACCTTACTAATAGTAAAATTTTTACAAGCATTAGGTTATT  
4509 TTCTTTGATTCATCAAGTAAATATTAATACTGTTTGAACATGTGATAGTCCCAGCGACTAGATTTGTAAAATATTTGCAGGATCAATGAT

FIG. 12B  
1 of 10

4601 TTGGTTTGGCAGAAGTAGGTAATTTCTAAAATTAAAAATGCAGGTAAAACAGGGACTGGAGAGGAGTATTTTTTCCTAGTGATTAATAAAC  
 4693 CTTTATTTTTCTTATTGTTTTGTTGCTTACCCAGTTTATTTGGCGTAAATCTGAGAACTTACTTTTTCCATGAGCAAAGTTAGAGGTAAAC  
 4785 TTTAACAAGCAGTTAGACAGAGGTAATGACCTTTAGATTAAAAGGTTTTAGGTCAAGCTGTATAAGTTGACTTGTGCTTAAGACATGATGA  
 4877 GCCTCTGTTTAACTGAAAGTCAAGCCCAGGACGCCTGCCTTTTCCATCAAAGACATGGGATTGGGTGGCAGCTGACTATTGATTTCCAATG  
 4969 ACGATTCTTCTTCAAGTGGAGGTCTTTTACCAGATGGTCTGTTGGTGGGGACATTGTTAACCCTGCGATTAAACCGACGGCATCTTCATCT  
 5061 GGCTTTTAAGCTCCTTGTATCCTGACTTGTACACAGCTTACTTATGCTTGTGCGACTATGTAAAGTGACAGTATATGAGAAAGGTAGTGAG  
 5153 TAGTAAGAATGTTGGGAGACAATTTAAGCTACCATTCATATTTTATAAAAATTAGACTTTTGTGTCTGGTGTAAACAAACAGAGGACAGAGC  
 5245 TTGTATGAAAGGATAAAAGAGCGTTAAGGGTTACACGTCCATTAGGATAAAAAAACTAGAATATTTCTTTCTGAAACCTGAAGCCCAGGCCG  
 5337 GGCATGGTGGCTCACGCCTGTAATCTCAGCACTTTGGGAGGTTGAGATGGGAGATTGTTTGAGCCCAGGAGTTTGAGACCAGCCTGGGCAAC  
 5429 ATGGTGAAACCCCATCTCTATTTAAAGAATAAGGCTGGGTGTGGTGGCTCACACCTGTAATCCTAGTGCTTTGGGAGTGTGAGGCAGGTGGA  
 5521 TTGCTTGAGTTCAGGAGTTTGAGACCAGCCTGGGCAACATGGTGAAACCCCATCTGTACTAAAAATACAAAAATTAGGCGGGTGTGGTGGCG  
 5613 CCCGCTGTAGTCCCAGCTACTCAGGAGGCTGAAGCATGACAATCACTTGAACCTGGGAGGCAGAGGTTGCAGTGAGCCGAGATCATGCCAC  
 5705 TGCACTCCAGCCTGGGTGACAGAGAGAGACTCCGTCTCAAAAAAATTAAAAAATTAGGCTGGGCGCAGTGGCTCACGCCTGTAATCCCAGCA  
 5797 CTTTGGGAGGCCGAGGTGGGCAGATCACGAGGTCAGGAGATTGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAATACAA  
 5889 AAAATTAGTTGGGCATGGTGGCAGGCGCCTGTAGTCCCAGCTGCTCGGGAGGCTGAGGCAGAAGAATGGCGTGAACCCGGGAGGCGGAGCTG  
 5981 GCAGTGAGCTGAGATTGCGCCACTGCACTCCAGACTGGGCGACAGAGCGAGACCTGTCTCAAAAAATAAAATAAAATAAAATAAAAAA  
 6073 ATTA AAAAGAAAAGAAAAGGAAACCTTAAGCCTAGTTATTGAGGTAGACAGGATGCTACCCCTGCCCTGTCATTTTATTTAAAGAAGCAT  
 6165 TTAAGCCTAATGAACACGAGCAGTTCTAATGTCCGTTGGAGGGGAGGTAGCATTACAGTTTATAGATTCAATTTAGCAATTACTGATTGAGC  
 6257 ATCTTCTGTGTGTCTAGTTATCTATGCTCTTAGGCGCTGGGGATGTGGCAGTGAACAAGAACAGATGTAAATGACAAGAGATGGATGGTGGT  
 6349 GATGGTTGCACAATTGTGTGAATGTACTTAATGCCACTGAACTGTATACTTAAAAATGTTCAAAATGGCTGGGCATGGTGGCTCACGCCTGT  
 6445 AATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACCTGAGATTAGGAGTTCGAGACCAGCCTGATCAACATGGAGAAACCCCGTCTCTA  
 6533 CTAAAAATACAAAATTAACCGAGCGTGGTGGCGCATGCCTGTAATCCCAGCTACTCGGGAGGCTGAGACAGGAGAATCGCTTGAACCCGGGT  
 6625 GGGCGAGGTTGCAGTGAGCCGAGATCCCGCCATTGCACTCCAGCCTGGGCAACAGAGCAAGACTCCATCTCAAAAAAAAAGTTTCAAATGG  
 6717 TAAATTTATGCATATTTTACCAGAATAAAAAAAGGCAGTTAAGACAAGTAAGATGCTGTGTGTCATGGGGCTAGATCAAGCACTTAGGGGTGGG  
 6809 GTGTTAGGGACTTTGAACGGGGCCTCTACCCCTGCGGGAAGGGCTGAGCTGGAGGGATCTGTGGGCCCTGATCAAGAAAGAAGCAGGAGCTG  
 6901 TAACCCAGCCTGGCTTTGAACTTGAGGCTGCCAGTGGAATCTGTTGTGTGTGACGGAAGGAGGCAGCTGCACTGGATGGGAGAACTGGA  
 6993 GGGACTCTGTGGCTGCCAGGGCCAGCTGCAGGGCACACAGCTGCACTCTGAGGCTGGCACCTGCCTCCTTCACTTACCCAAGGCTTTTCCTC  
 7085 CTGTTTTTGTTCAGACGGCCATCCTGAGACGACAGGGTCGATACATATGCTCAACAGTGCCCTGCGAAGGCGGAAGAGGAAGCACAGAGCT  
 7177 TGTGTTTACAGAGGTAAGGCTCTTTCCTGCATTAATTTACATTTTGAAGTCATTTTCCCTAACTGCCTCCTTTTCTTTAAATTTCAAAT  
 7269 TGTCAAGGAAGTGTCAAAGGGTAATTGTATTTCTATGATGGAAGTTCAAATAGAATAATGTGAATTTTTCAGACTCTGAACTTGACAGA  
 7361 AATGTCCACAGGGCTATTTCTTTTTTACATTTTTTATTATTTTTTAAACTTTATTTATTTGGAGGGGGCTATATCTGACTACAAAAAGTGAA  
 7453 TTCCACAGAATTTATCTCATGGACTTAAAATAAGCAGTAACCTGTAAATGAATTCAGTGGAAATCTGTGGGAGGTCTTGTATTGATACTGT  
 7545 TTTTAAGGGTGACACACACATTTATGTATCATTTATTTTCACTTATATATTTGCAAGTTATTTTTGACTAGTATTATGAAGTACTA  
 7637 GCTGATAATAAGCAGGGTCTATCGCTAGTCAATATATATTTATTATATATATTGATTACTATATATATTCCTAATCAAGATACATTGATTAAT  
 7729 ATTATTTTTGTTTTGAAAATGCAAATAAAATTATCTTATGGAAGAAAGATAAATTATTTACTTTTTTATTTTTTATTTTTTATTTTTGAGACAG  
 7821 AGTCTTGCTCTGTTGTCTAAGCTAGAGTGCTGTGGAGCAATCTTGGCTCACTGCAACCTCTGCCTCTCCTGGGTTCAAGTGGTTCTCCTGCC  
 7913 TCAGCCTCCCAAGTAGCTGGGATTACAGGCGTGCACCACCACGCCTGGCTAACCTTTGTATTTTTAGTAGGGACAGGGTTTCAGCCTGTAG  
 8005 TCAGGCTGGTCTCAGACTCCTGACCTCAAGTGATCTGCCCCGCTTGGCCTCCCAAAGTGCTAGGATTACTGGCATGAGCCACTGTGCCTGGC  
 8097 CAGAAAGAGAAATTATTACAATTTAGGTTGTTTGCTTTAGTTTTTCCCCTTGGAGTGTGTTTTTCCCTCCAGGTAATTTTAGGTAGGAAG  
 8189 GAATAATTAGATGTTTTAATTTTGTTCCTTTAAGTGCACCTTCATTTAAAAATAATGATTTTTTTTTTAATCCTGGTTTTCTAGTTGATATT  
 8281 TAGATCATAAATATGCTCATCAATAAAATTGCTTACTATAAGGAAGCTATAAATACCTTATAAAGACCAATTAAATAACAATATTTAATTT  
 8373 CCATTGAGATTTTTGAAAAATTAAATATAAAATTAAAAAATTTTAAAGTGTGTTCCCTCATCTTTCTGAAGAAGTAACCTCCTGTCTTACCT  
 8465 CCTTTGCCACTATATTAGTAACTTAATTCAGACAAATACAGCCAGATATGTTTGTGAATGTAGTTATAAATGTCCTTTTTAAGGCAGGTAG  
 8557 TGGCAAACTATGACCTGCAAGCAAAATCCAGCCTGTAGCCAATTTTTGTAAATAAAGTTTTATTGTAAAGCAGCCAAGCACATTTGTTTACA  
 8649 TATTGCCATATGGCTACTGTCACCATGCAACTCAAAGTTAAGTAGAGATAATATGACCTCAAAGCTGAAAATATTTATGATCTTGCCTTTTA  
 8741 CACAAAAGTTTGTGACCTATGTTTTAAAGCATGTGGCAAAATTATTAATTGCTAACTCAGTTCTCCAGTTGATTAAAAAATATGGTTT  
 8833 TTTGAGGGAGAACTCTCCATTAAGTTATTTAATCACTGCAGGTTGAGCAATAGCTGCTTCATCCTATGCTGCTGGAGCCAACATAACTAAAC  
 8925 ACTTTTGGGACCCTTCCACTTGGGTGGAGTGAACATCACTTCCTCTTCATCCTCTGATCCAGGGAATGACCTAATGGCTTAAACAAAGCAA  
 9017 AACAAAGCAGAAAAAACTTCAAAAACCTTTCAAGTGTAACTTCAAAATATTATTGAATTTACCAAGTTTGAATGTAAATCGTATAATCAGT  
 9109 CAGTCACATTCCCTGTCTTTTTGAAGGTACAGTTCTCAGGATCTGGCTTTCTGATGGAACCTTATCTCCTAGACTTTCTGCATCCCCAGAG



9201 GGGTGGGGGTTGCTTGCCTAATTCTGTGCTTCTGCTTTGGAATTTAGCCAATGCCTGCTATGTAGATGGTCAACACTGGCTTGTAAATCAA  
 9293 TGAATCTCTAAATACTTAGCCAGGTTCACTGTGGAGTTTTTTTTTTTGTATGTGCGATTCTTTATAATTATAAATATA  
 9385 TGAACATAATAAACTGTAATTATTTTCCCATTTTGCCTAAGGTATTAAATCTTGGCCAGGCGCAGTGGCTCACGCCTGTAATCCCAGTAC  
 9477 TTTGGGAGGCCGAGGCGGTGGATCACGAGGTCAGGAGTTCGAGACCAGCCTGACCAACATTGTGAAACCCTGTCTCTACTAAAAATACAAA  
 9569 AAGTAGCCAGGCGTCGTGGTGTTCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGTAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGC  
 9661 AGTGAGCCAAGATTGCACCACTGCACTCCAGCCTGGGTAACAGAGCAAGATTCCGTCTTAAAAAAAATAAATAAATAAATAAATAAATCTGGG  
 9753 CTCCATTTCATAAATGCATTTATAATATACAAATTTATGCTGGTTTATTTAGTTTTTAATTGGGCAGGTACATTTTATACTTCAGAATAGTTT  
 9845 TAATGTCTTGCAGGAGACTTGGTGGTACAATTTCAAGTTCTAGAGCTTGTGAAAAACCGCTGTAGCTCTTTGAAGAGAATAGGATCACACCC  
 9937 ACAAGACCACGTCGTACATAGTCAGGCTGACTACTTAGCTGTGCCAGTGGACCTAGGGGCGAGTAGTGGGAAAATCCAAACTATGGATTAT  
 10029 TCTGAAATCTGACTGGTTTTGGAATGAAATTTTGAAGTGTGAATGTATGATACTCAAATGTGAAAATGCTTTGAGAAATCTAAAAAATCAT  
 10121 TTTCAAGGATAAAATAGCATTTTAAATATTTCTTAACACAGAGTAAAAAATCAAAAAGTTTGCCAGCTCCAAAATCTGACTTTGCGCTTA  
 10213 GCCCTTCTCTGCCATTTATTTCCACCTCTCAGAATAATTTTAAAACTTACATTACCTTCACATCACACATCACACCCTTGCCAGTAAAGTGC  
 10305 TTATTGTAGAGCCTGGCACAAAATAAGCACTTGCTGCTGTTGTAATTTCTACTTTTCATGAGTTTGGAGGTGGAGGGGAGTTTGCTTCAGAAG  
 10397 ACCTTCAGGCATGGCCTGTGACACAGATAGAAAGGTTGACAGTAGAAACGATGAAGTGAAAGAAGCAGAATAGAGACGTTGTTTCATGTTAGT  
 10489 TACAGCCAGAGCCAAATAAACTAGATAAGGCATAAATACATTTATCGACCTATGTTTTTCATTGATTCATCTCATGCTTTTTTCCACCCACAC  
 10581 TGAAGTCTGCACATAAATAGTCAAGTGCTGACTGGACAGAAGATCATTTGAACCTAAGCTGATTTTAACTAAAATTTCTTAGCTTTTGATTTT  
 10673 GACTCATTTCATATTTGAAGATCTGGTGGATGTGGCATTATAATCTTAGGCTTTATGCATATTTGGAGGATCTTCATGGGTGGATGAACGACG  
 10765 GACACCAGGCTGGCCCATTTGAATCTGTCTATGTGGCGAGGTAGTTTGAATGCAGTCTGTGTAGGTGAGGTCACTAATAGAATCCTTGCCAGTG  
 10857 GAAGACTACATTAAACATTTCTGTACACACACACAGATGAAAATGTCATGGTGCTAATATTAATTGGAATGCATATGTTTCTTCTAATTATT  
 10949 TCACTCTTTTCTTATGAATGGAAAGAAAACTGAGGGCAGATTTATATGGTAATATTAACCTAACAGAGCTAGCTAGTTCTAATATTAGTAA  
 11041 ATACCATGAAGGCGTCTTTAAGTTTACGCATCCTATTCACTTTTATGTCATAGTAGAGGCCACATGGCTTTTTAAGAACCAGGTGCTTACTC  
 11133 TGAAATAGTCCATTGTTAAGTAAATGTCACAAGGTTAGGTGAAGTTGTCTCCTTGTAACCTGCTCTCAGATCATTAATGATGATAACTTA  
 11225 AAGTGATACTACCCCAAGGTAATGTTTCAAGTCTCAAGCTTCACAGGAATCTCTGGTGGTGAATTTGAGGCATTAGCTCG  
 11317 TTGATGGGAAAAAATTTTTTTGATGTTTCTATGGGTATGCCTTCACAACTTTCCCAAGTGTTTCCAAAAATCACGCTTTTCTGTTTTTA  
 11409 TTTTCTATTTTTTAAATCCCCCTCAATAGTGCATCAAAACCTATAACTCTGACTGGCATCTTGTGAACCTATAAATATGAAGATTACTCAGG  
 11501 AGAGTTTTCGACAGCTTCCGAAGTGAGTAAGCTATATTATACACATAGGGAAAAGTCTTTGTACTTGAAATGCTTGGGGGGAGGTATGTAAC  
 11593 TCATATGCAATCAGAGTAATTGAGGAAAATATTTTATAGATGGTTTATGTGTATGTGGTGTAACCTATTGTTTACAGGGCCTTTGATTGTAAGA  
 11685 CTCAAACATGCTACTTTGGTATTGATAGGTAGTAATAAATCATGTGGATGGTTTAAATTATGTCCAGTGGTTCTTTTACAGGTACCACTGATAA  
 11777 ATAAATAGGTGAAATTTTTTCACTACAAAATATAAAACAAACAAACGTAAGTGACTGAAATATGCTGCAGTGCATTTTTTCTTGAAAAG  
 11869 AATATTTTGAAGAAGATAATTATAAAGAGCATTTTACATTGAATAAATTTTATGTTTTTAAAAAAGTAAATCAAGAAACAAGCATTTTCCC  
 11961 AGTTAAATTTTTTTTTTATCTCCTTTAATGTATTAACCTATTCTAGACTATACCAAAGCAAGTGATTTAGATTGAATAGTTGTGGCCAAA  
 12053 GTGAATCGCGGTAGCTAGGTATTGCCTTGACAGACTATCTTTTATAAAAGGTTCCATTGTGTGTGCTTTTAAAGGAATACCTGAGACTGGGT  
 12145 AATTTCTAAAGAAAAGAGATTTATTTATTTATTTTATTTATTTAGTTTGTGAGACAGAGTCTTGCTCTGTCACCCAGGCTGGAATACAATGA  
 12237 CACGATCTCTGCTCACTGCAACCTCCACCTCCCAAGTTCAAGCAACTCTCCCGCCTCAGCCTCCCGAGTTGCTGGGATTACAGCCACTTGCC  
 12329 ACCACGCCCAGGCTAATTTTGTATTTTATAGTAGAGACAGGGTTTACCCTGTTGGCCAGGCTGGTCTCGAACTCCTGAACCTCAGGTGATCCAC  
 12421 CTGCCTCTGCCTCCCAAGTGCTGGGATTACAAGCATTAGCCACCGAGACTAGCCAAGAAAAGAGATTTATTTGCTTATAGTTCTGCAGGT  
 12513 TGTACAAGAAGCATAGCACTGGCTTCTGCTCAGCTTCTGTTGAAGACTTTTGTGCTGCTTCAAAACATGATGGAAAAGGTCAAAGGGAAAGT  
 12605 TGGCACTTGTGAAAAGAGACCAAGAGGAGGAGGAACTCACTTTATAACAACCCATTCTCTTGGGTACTAATCCATTTGAGCAACAAGTAA  
 12697 TCCCATCTTGCCAGGAACAAGAATTCATCTACTGTGAGAACAGCACCAAGCCTCTCATGATGGATCTGCCCCATAACCCACTAGGCCCA  
 12789 ATTATGCAACAACGGGGACCAAATTTCAATCTGAGTTTGGTGGGAATAAAAACCATATCCAAACCATAGCAGTTGGCAAATTGAATTATAC  
 12881 TTGATTTTGGGAAAATTGAAAGCAAATAGTGATGGATTATGTTTTAAACTAACCATCACCAATGAAAATATTACTTGAGACCTGATTAATG  
 12973 TATTTTCTTTGCGTTTGTACATCTTTGAGCTGGAACCTTTTATGCGGTTCTCAGTAGACCTAGCTGTTTGTTTTTCTCCTTGTGTGGCT  
 13065 TTGCCACTCCTTAAGAATGTTTCGGCCAATTTCCCGATTGCCTCTTTTAAACCTCAGCCAGGAACACTCCCTCCTAGTATTATCTTCTCCAG  
 13157 ATGGGTAGCCCTTTAGTTCTATATTTACCCAATCTCCCTTAGGGATTTTTTAATTTCTTCCCATTTGGATTGGCTTAACCCATCTTTGTTGAT  
 13249 CCTCTTGTTCATAGTCTCAGGGTTGAAAGATATCAGACTATGTCATGTCGTATACTTACTATCTAATAGACTGCTGGTACATTTTCTCTCTT  
 13341 GGCATTAATGAGAATTTCCAAATGTGTGATGAGAAAGAGAGGGAGAATTGTAACAGTGGTGAAGAACACAGATTCTGTGTTCTGCCTCAGAC  
 13433 GACTCTGCCGTCTGTGAGCCTGATTTTCTCATCTGTTAAATGGCCTTAACAACAGTCAACCATGATTAAAGGATTAAATGAGAGGGCACATG  
 13525 AGAAGTGCGCAGGGCCTAGCACGTCATACCCATTGAGTAAATGTAAGCTGCTTATTGGTATTGGTGTCTTGTTTTTGTTGTGGTTAATACCA  
 13617 TCATTGTTAATCGGTTTCAACGCAACAACCTCAATCTTCTTTTTTCTCATAAACTTTGATTAAAGTTATTCTACCAAGTCTTTGTTTA  
 13709 TTAATAACTAATCCACTTTCTTATTTTATAGTATGCCTGCTAACTCCCCAGAAGCTATGCTGTCTTTTCCACATAGCTTTTTTGGAGCTTTCTT

FIG. 12B  
3 of 10



13801 ACTCAAGTCTCTTGGCTTACCCACCTTGAAAAGCAAGGGCATAGATGGTTTTATTCTTTGTCTGAATAAAGAAGCTGGGCCATCTTTGGATT  
13893 TAGTAAAGGCCGGGCCCTATGATGGAGGAAGAAATGCAAAGCCTCTTCCTTGACTAGGCATTTCTAAAACAAATTTCTTTGTTGTTTTAT  
13985 AGCAAAGTGGTCAAGTTGGATAAACTTCCAGTTCATGTCATGAAGTTGACGAGGAGGTGACAAAGATGAGGTGGGATACCTGCTTGCTGT  
14077 TGCTTCTCTTTTCACTCTAGATTTAAACATCAATTTTACAGACTTAGAAGATTAGTTAGAAAATTACCGACATTTAGCCAAAACAGGCATTG  
14169 GAGTGTTACATGAAACGGGAATAATTTTTTAAAAATGTTATTGATTGATTGGAATAAGGTCTCTGTTTCAACTTTACTGCTTAGCATTTCAT  
14261 GTTCTCTTGGTTGTGTTTATTTGTTCTGAGATCATTTTCAAAGACTTGGATCAGATCTGGCTACATTGTTAAAAGATATCAAGATGACTTAG  
14353 ACCTTGAATTTAGGTTGTTTTTCAACAGATCTCGAAACAGCTGCCAGCCAGTAGATTTAAATGGCTATTTCTTCAATGATTGCTTTTAGTGA  
14445 AGTCTGATTTGATCAAGCCCACTCCCCCTATTCCCTAGAGGAAAGCTCATGGCTAAAGAACTATATAAAGGGAGTAGGGCATTGAGATGAGTC  
14537 TGCCCACTGAGTGAGGGAAACCTCACAGAAGACAATGCCCATCTCTGCATTTCTCATCCTCCCCATTGATTGTTAAGTGTCCCATTTGTGAG  
14629 TTTAGGTTTTTTCCTTCTTTAAAAAATTGTCAGCTGAGCTATAACATTAGCCACTCATTAAGCAATGTGCATGTAGCAAATTATTTTTATTC  
14721 CCCCCATCACTTTATCTCTCCTTTCTGTATTGCCTCAATTTCCCTCCCTTGCTTTATTACCTTTCCCTGAACCTAAGCCTCTGGGAAGGTTTC  
14813 CAGGAATGTGCATGTGCTTTTGTCTCTGACTATAGGGGAGTGTCATTTGAAAACATTTTTTCGTGAACCAGGCAAGACCTTCCAACGTGA  
14905 GTGGTCAGTTGAGGTATGTCCTTTTTTGGTCTTTTTGTGGCTCATTAAACACTGACAAATAAAAATTTGGACAGGAGCTAGCTTTGCCTTTAA  
14997 TGGAATAAAGTTTTTCAGAAATGTAGGCGGGTCTCTCTCTTTCCACCGCTAAGTGGACTTTTATGTGACTTGTAGGCATTGGTGTCAATGGGTG  
15089 CTTCAAGTAAAGGGGCAATGGACAACCTTGGCACAAAGGGAATGACCTTCCCATTGACCAAACCTCACAGCAAGCAACCCAGGTAATAACGGGAG  
15181 GTTGAAGTGTGGTGTTCCTTCTCTCTAGGATGCTGCCTCCCTTGGTTCCAGAAAGGGTGGGATCACCAAGCATGGCTGGCTGTACAAAG  
15273 GCAACATGAACAGTGCCATCAGCGTGACCATGAGGGTGAGGACGCACATCACTTTGCCCTCCCTCTCACAAAGCCCTTTCTGCCATAGAGCT  
15365 CGAGAACAATGCTCAAGATGAATGCGCATGCTGTTCTTCCCCACAAAAGGGACATTGTCTGATTCTTAGGATGCTCCCCTGGTGATAGCACC  
15457 CCCATTGGCACAGCCTCATCCACCCACTTTCCCTCACTGTCTTCTGACCACCAGCATAAGGAGACCATCCCTGGGCTGGTGTGAAGGTGCAG  
15549 ACACTGACATAGGCTTTCTTCTCTGTAATAACTGAAAAGTGCTCTTTGGTACCTCACAGAATGTACCAAGGGGCTATCTGTCTATGCCAATC  
15641 CTGAGCACTTCTGTGGAGGTGACTGCAGCAAAGTCAAGTAAAGCAAAAATTGAGGACGAGAAAAGAAAATAGTTGCATAGAAGAGAAGGTT  
15733 GCAGACAGAGAAGTCAAACCAATAGAAGAAGCTATTCAAGGAGAAAAGTGGGACCAGAGGAACATCAGGATTAATAACAAAGGGAAGAGAAAC  
15825 AAGGGAGTCAGGGAGATAAAAATTAAGGAGGAAATGTGACTGTCATTACCCTAAGGCTGGAAAATCATTCAAGCGTCATGAGGCAAAAATAG  
15917 TTCCCATTCTGTGAGCAAGAAACCCTGGGGATTTTAGAGAAAGTTTCTGTCTTCTGTGCTGCATCCCAAATTGGAAGTCCCTGCACTGCTT  
16009 TTGGGTAGTTATGTAAATCTCTGATTCCGTGGGTGAGAAAATGACCCATGGATATTAGGGGAACCACTCCTCAGAACTGAGATGCAGTG  
16101 AGCTTCTTAGATGGGATGGGGAGTCTTGACCCACAGTGACCTGGAGCATCAGCTAGAGTGAGAACGGAAACAGGTTTTATGTATGTATGTA  
16193 GTCATAAGTGGGTATTGATAGAGATTGTGACCCTCTTCATTTTGAAAGAATAGCTGTGTGTATATTTTTCTCTCAGTCATTTAAGAGACGA  
16285 TTTTTCACCTGATTCAACTTGGCGATGGATCCTATAATTTGAATTTTATAAAGATGAAAAGATCTCCAAAGAACCAAAAGGATCAATATT  
16377 TCTGGATTCTGTATGGGTGTCGTTCAAGTAAATATGAAAAGAGTTTTACCATTATGTTTTCTTATCTGCAGTAGTGCTTATGTGTAAATTA  
16469 GCAGATTTAAGCAAACACTTCCAAAAATGGCAATATGCATGGTAGAAATATAACATATAACTTTAAATGAGGCAAGCCTTGTTTTTTCATCAT  
16561 TGTAGAAGATGGAAGGGATAATGTAGAGGCAGAAATATGCTGTGGCAGGCAGGAGCACTCTGGCTCGGCCACTTTATAGCTGCGTGACCTTT  
16653 AACAGGCTACTTAATTCAGATAATGAGAATGTTTCTTTAATACGGCAAATGAGTACATTGGATGAATCAGTGCAGGAAAATATTTAAACAC  
16745 TTCATAGTATCTCAGTGGTGATTTTTATCGCTAGCATTGTAGTACCAGTGGCGGTGTAGATCAGTAAAGAGATTAGGTTTCAGCGCAGATTG  
16837 AGTTCAAATCCCTGCTCCTCCACTTACCAACTGTGTAACCTTGGAGATGTTATTTAACCTCTCTGTACCTCAGTTTCTTCATTTGTTAAATA  
16929 AGGATAATGGCAGTACCAAATATGGTTACTGAGAGGGTTCATTCAATACACATGTAAAAAGCTTAGAACAGTGCCAAACAAATGGTAAGCATT  
17021 TGGTCAGTATTAGATAGTTTTGTTATCATAGGGCTGTTGTACTTTTATATCATAGGGCTTATGTACTTATCCTTTAAAAATTATTGTTAATTA  
17113 AAGATAACACATGAATGTATTTTTCTTGTAAAAAATCAGCCAATACAGATAAAGTGAAAGTCCTTCTGGACTCCTCCCCCTCCTTCAGTGTCT  
17205 CTTTTCTGAGGGGAGCTACTACCAGTTTTGCATGCATCCTTCTGTAGCTTTTTAGCATTGTCTTTGGAAGAGAGTTGTCAATTTCCCTGTCC  
17297 ATCATCTGTCCATCCATCCATCCATCCATCTGTCCACCCCTCCATTCACTCCAGCCTTGCCACTTTCAAGGAAGATTTAAGGCAGCAGC  
17389 TTATAAGCATAACACAGGACATGGGATAGCATAAATTTAAAGTGGGGGTGAAAGCAGAAAGATGAACAGGGGATTGGGATAAGGGTGAGAGAA  
17481 AATAGAGTTAAGGAGAAAGCGTATGTTTTGAAGATCTAACACCTGCTGTGGGTGGGCCACCACCTGGGCTCTATGCTTTCTCACTTGGAGAC  
17573 CTGTTTAGTCACGCAATTCACAGTGCACATGAGATAAAGGCATGATGCCTGTTTAGTCGACTCTAGAAGCACCCCTGACTTTAAAAAGAAGT  
17665 TAAAGCAAACTAAATGTATTTGGCAACCTCATTTTTTAAAGTAGGAAGTAATATTTTTGTTTTATAAGAGAGTTTTGCTGCCTGTTTCTG  
17757 GCCCAGGGACAGATGTTTATAAGTACAACAGCCCTGAGCTATCAATTAGTCTCCGGGGTGCAATTTCAAATCTAAGGTTCTGACTTCAATGG  
17849 AAGTCTCTTCCTTCAAATTGTCTTTGCAGATGCAGCTGATGGTGTTCATTTAATAAAGTGTATCCAAGGCTTCAAAAAAGTAAATAATT  
17941 TGTTTTTATCTGTGTCTGTTTGTAAACTAAGCATCAAAGTTGTGATTTAAGTGTTTTTTAAAAATTATTACTTATGGATATTATAAAAAAAT  
18033 TAGTTGACTGGTGCTGTGAATTAATAAAGTGCCATAACTAAAAAATTTTGAAGCATTTTAGAACCTTGAAATTTATTATACTTATTTTGC  
18125 AGATGAGAAAAGTGGCTCAGAAACAGAAATTTAGAATTGAGGCCTAATGTTTTTCTCCACTTTTAACTTTCTCTTTTCATGATTGTGA  
18217 GTATGCAGGGAAAGGAGGAGAGAAATTCATTTTGTTCAGGCCTTTGACTTCTTCCCTGGTTCCTTGCCCTGAAGTTTAAAGTGAATCCAA  
18309 GTGGCAATTACTGAGCCACAGCAGACAGTCTGTGCACAAGAGTGTGTGGCTTTGCCAAGGGGAGCACTTGACTTTGCATTTCTAAGAACTG

FIG. 12B  
4 of 10

18401 TGCTGCAGAATCACAGAGACTTTTGGGAGGGTTGCCCTGTCCCTGAGACCTCCACCAAGGAACCTCTTAGAGAGAGTGTGGATAACCCAGTAG  
18493 GATTTTAGTGGCTATGCGGGGGGCTGTCCTGCTGGCTCAGGTTAGTGGGAGTGTGTTGATTTTCATATCGCTCAGCCTGTCTTACAGGGGATC  
18585 TTGTGCCATGATCCTCAGAGCTGAACCTCTGTCTACTGCGGCCAACCTGGGGAGATTTTGCTCCCTGGAGGACATCTTGGAATGTCTGAAGA  
18677 CTGGCATCTATTGGCTTAAGGCCATAAATTTGCTAAACATTGTACAATGCATGGACCAGCCACTCACAACAAGAATTGGCTGCCCAAGTG  
18769 TCAGTAGTACCGAGATTGAGAAATCCTGGCCTAGTGCATGTTTCATCTTCCGTCTGTTACTGCACATGGACTACTGTTCTTGTCTGTGAGCC  
18861 AGTCACCCCTCTTGCAGGCATGAAAACCTGGAGGCATGAGGCAAGGCCACGGACAGGGAGTCCAAATACCTTTTGGGATTTCATAAAGGATGGGA  
18953 AAGTTCAGATAAGTAAGCCAAACATAGTAATAGATAATGGTTGGCTTTTAAAAATGTAATACCATACTACTTCATTAAAAAATAGGAG  
19045 CTGAAGAAATATGAAAATTTTACATGAAATTTTCAATTTATCAACAAATATTTTCAAATACCCACTATGTGCAAGTCACTGTAGAGTCCATA  
19137 GAGACTAAGGATGTGTAGCACTGACAAAATGGGAGCACTGAGGAGGTTTCATTCCTGTCAGGGACACACAGTGAATCAGATGAGTATGTA  
19229 AAGCAGGTAATGAGTCAGAAGGAAAAATAAGCTTGCAGAAAGTGAAGCAGGGAAGGTGGACGGTAATGGGATTTCATGGGGGGGGGCTTTC  
19321 ATGAGGAGGGGGCAAGCTATTTAAAAATAGCTTGGTTCTAAATGCCAATGAGATATCACTCACCAACAAGAGAGAGTAATTATTTTAAAGCAG  
19413 TTCTAATTCTTTTAAAGTATGTCTGTTTATCCTTTTTTTCATTTTCAGAACAAACAAAGTCAGGCGTTTTGCTTTTGAGCTCAAGATGCAGGACA  
19505 AAAGTAGTTATCTCTTGGCAGCAGACAGTGAAGTGGAAATGGAAGAATGGATCACAATTCTAAATAAGATCCTCCAGCTCAACTTTGAAGCT  
19597 GCAATGCAAGAAAAGCGAAATGGCGACTCTCACGAAGGTAGATAGGCTTGGCTTCCCCAGGCACATACACACTCTGTGGGTGTCTTTATTT  
19689 TTGCCAGGTGGGTATAAGAAGGAGACCTGTGTTACACAAGTACATGAGAGGTGGGACGGATAGGAGCTCTTTACAAATATCCTGTCAGCAAA  
19781 GGTTTTGTACATTATAACTTACTTCCCTGACATTTCTGATATGGAATCATGTAATGGGAAGAACCAGCTTTGGAGGCAGAAAGGGAGA  
19873 CCTGGGTTTGAGTGCCATAAATACTGTATTTTCACTGTGTAGCCCTGGGTAAACAACCTTATGTTTTCTGAGCCTCAGTTGACTCACCTATAA  
19965 AATGGGAATAAACATGAAAATTGCTGGGAAGATGGGAAGTGTAAATAAGAAAATGAATCTCAAGTATCTGGCATAGAATTTTACTGTATTAT  
20057 AAAATATTAGTAATAATTAGAATGCATGGGAGCCTCAGATTAAATTGGTGAGAAAAATCTGGCTATGTTCTTGACAATTCATGTTTTACTTC  
20149 AACCTTAGGTGATTCCCAACCCTGGCTTCCCCTTAGAAGTACCTGGGAGCTTTTTTAAAAATACCATTTACCTGGTCCCACAAAAGATTCTG  
20241 ATTTAGTTGGTCTGGGGTGGAGCCTGGGCAGGTCTGACTTTTAGGGGTCTCATGGACGTGTCCATGTGGGCTGTTGTTTCATAGCTAGTGTC  
20333 AGTTCTAATTGGACGGTGTCCATGCTATACCAGCTGCTCAGTGTGTTTACCTTTCATCACTGAGCCTGTGGATCAGTATTTTTTCAAAGCACC  
20425 CCAAGTGTGTTCCAGGAGCATCCAGAGTGGGAACCACTGTGTTCAATTTGAAGGCACCTAAGAGAAACGGCCTTCTCCTCCTGTTTCAAAT  
20517 GAAATGCTATGAATTACAAGTGAATCCGATAATCTGTTTTGCCATTTTAGATGATGAACAAAGCAAATTGGAAGGTTCTGGTTCCGGTTTAG  
20609 ATAGCTACCTGCCGGAACCTGCCAAGGTAACATCGTCTTATATCTTCTGCTCTTCTGTTGAATGCTGTTGAAGTATGTCTCATTTCACTGGTT  
20701 TGTCCAGAATGGAATCTGTTGAAATCATAAAAATTACATTGTGATTACCTCTCTCTTTTTCTGACCTGATTACGAGGTGACGTGTACTCATG  
20793 CAGTATGATTTTCAAGTCTGTCTTCTAAAAAGTACCCTACAAAGCATTCTCCTTTTATTATTATTTTAAAGTGTTTTTTCCCTGATAATGCTT  
20885 AACACTGCATCACAGGTAAGTGAAGAAATAACTGAAATATGCAGGCAGATGTTCTCATAATAGCATCGTACTTTCTATGTTGATACATGTGCT  
20977 CTCCCTTACTCAGGGTAATAGACACGGTTCCAAAGAGGAAGGACCTGGTAATCTTGCCACGAAACCCGGGGGTGCTGAGTTACAGAAATT  
21069 GTTTTCGGGTCACTCTTACTGGAAAAAAATAAGCTATTCCTGTGTCTTACAATTTTGAGAAATTTAAAAGTTACTGAAAAGCACAAAGAAA  
21161 GCAAATCAACCATACTGCTACTTCCAGATTAAATATCTATTATGATGTTGCCCTTTTTAGCTTCCATATTCTTAAAAGATATAAAACATCGG  
21253 TTATAGTTGAAGTTCTTTTTAAACACTGTCTTATTCCCATTTCTCTTCTTCTCCACAAACCCCAATCAGAAACAAGCACTATTAAAAGTTTT  
21345 ACTTTTCATTTTTATATCTTTACAAATAAATCTATCATAATGATATATACAGTATGATTTAGTATGTTTAAAAATGTTTTATAAATGCTAACA  
21437 TACCATATGTATTCTGCACCTTAAAAAATTTTTAAATTTACCCCTTTTATTTGTACTATATAGACTTTTTATTTTAGCTGTTCTATTATTTT  
21529 CATTTTTTCTATTATAAACAAGCTACAATGACTGTCCTTGTACTTGTGTCTTGTGTGATCTGAATGATTCTTTCTTAAATGAGAGAAA  
21621 TATCTTTGTCTTCTAGCCCTTTGCACTCTTACTCTGTTACTGCCCTTCTATTCTTTTTTGATACTAGAGTGAAATGGCGACCCTCCACAC  
21713 CCACATCTTAAACACTATAATAGAAACATGGTTTATCTATATAGGATTATAAATAGACCAGCATTGAGCATTGACCTTTATTTTAAAGACAAC  
21805 ATGGCTGTTCTCAAGTGTAATAATCTCCCTCCCTGGCTAGGGCTTTAGAGCATTGTTTTTCTTTAGGACTTGACTGCTACCACAGTATCTTTT  
21897 TAGCACCTGCCTATTAAAGCTAATTTTAGTGCCACCATTGTAAACCACCTCCTAGTCTGGGAAGAGTTTTGGCTTGTGTGTTTGTGTTATGA  
21989 ATGCTGTGTATCATATTTTGCATTGAGATTGCTTTTTTGTGTTCTGGATGTTTGGGGGTTTCATAATTTCTCAAACAAAATATTTGTGCCC  
22081 ATTTGGGTTTTAGTTTGTGTCAGCAGGTAATATATGTGATGCCATCTAGAATTCAGAAAGTAACCTTCTGCACTTACTGGGTGAACGGAATG  
22173 GATACCTAGGAGAAGATTCATGTTATTTGAGCCTAATGTTGATTAATTAATAATCTATGCTTTTTCTATGAGGATATACAGGAACGGTCCC  
22265 CTCCTTCTCTTACTACCCAGCCAATATAATTCAGTATTGTTTGATCCCAAGACCTAGGGAGATTTTTTAAAGATATACATATATTTAATATAA  
22357 ATGTATACATTTATGTATATATACTTTTTTATAAGTATAAATGTATATTTTGTCAATTTAAAATATTTGGAACATTTTTAACTATGGGTTA  
22449 CAAGTTAGGTTAAGCCATTTTAGTTGGTGAAATCAGTTTGATTTCAACCCCTGCTTCTTTTTGTTTTGTTTCTATTTAGTTTTTTATTCTTT  
22541 TTATTGAGGTATAATTGCAATAGCAGAATGCTCAAACATGAATTGTAGAGCTCACTGGAGTTCGCATTTGTACACTGATATAAGCAGCCCT  
22633 CAGGCTAGCTTGTACCTGAGACCCTCTTTATTTTACCTCCATCACCGTAGATTAGTTTTGACTTTTCTAGACCTTCTGTGAATGGACTTA  
22725 TACATGTACTCTTTGTGTCAGGCTTATTTAGCTAAACATGTGATTCATTTAAGAAGTTTTTTTTAGGTTCGGGCATAGTGGCTCATGCCTGT  
22817 AATCCCAGCACTTTGGGAGGCTGAGGTGAGCGGATCTTTTGAGGTAGGAGTTCAAGACCAGCCTTGCCAACATGGTATAAAACCCTGTCTC  
22909 TACTAGAAATACAAAATTAGCTAGGCGTGGTGGCAGGTGCCGTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCATTGAACCTG

FIG. 12B  
5 of 10

23001 GAAGGCAGAGATTGCAGTGAGCTGAGATCATGCCACTGCACTCTAGCCTGGGTGACAGAGCAAGACTGTGTCTCAAAAAAAAAAAAAAGGGTC  
 23093 CGTTTTAATGAAATAAAATGGAATGGAGAATATGAAAGTACACTGCCCTTAATAATGACATTATTTTTTATATAAAATACTGTCATTATTAT  
 23185 TTTGGTGGCACCTGCCACCATGCCCTAGCTAATTTTTGTATTCTTAGTAGAGACAGGGTTTTATACCATGTTGGCGAGGCTGGTCTTGAACCTC  
 23277 CTAACCTCAAAAGATCCACCCACCTCAGCCTCCCAAAGTGCTGAGATTACAGGCATGAGCCACTACGCCCACCTGAAAAAAAAACTTTTTAA  
 23369 AGTGAATTACATAATTTTTTACATAAAATAATGTCATTATTAAGGGCAATGTACTATTTATACATATAGTGTGTATGTGTCTTGCATAGT  
 23461 GATATAAAAGATATTTGTTTTCTTAGTGTGCTATTATGTATATTTATTTACTTTTCATTGGTATATAATGTACCTATTTTGGGAGTTCATGT  
 23553 GATACTTTGATATCTGTATACAATGTGTGATGATCAAATCAGGATAATTGGGATAGCCATCACCTCAAACATTTATCTTTGTGTTGGGAATT  
 23645 TGAAACATTTCTTACCAGGAGTCATGGTCAAAACCTGAAAATGAATCCTTGTTAGAGGCTTTTACTCTTTCCCCCTGGCTTTCAGGTGTTT  
 23737 TACAAATACTTTTATTTAGGAAGGTAGAAAGGTGGAAAGTAATTTTTTGAAGGGGAAAAGAATGAAGAAAATGGAGATGAGTTATTCACTCA  
 23829 GCACATGGGTATCTGTGGGCTTTGCCCTTTTAAAGCCCAGCTTGGTGTGAGTGTGAGCAGCCAGGCAGTAAGGGGAGACCTGTGTTCCCCAT  
 23921 CCCCAGCCTTGAGCAAAAATGCAGTTTTGGCTGTTTATCATCCCCCTTCAGGGTGTCTGAACTATTTGCACCGGTTGAGAAGGCCAAAGAGT  
 24013 TGACCTGATAACTGTTGGTCATCCCATAGGAAGGATGGATTCCATGGTTACAGAATCAGAGACTGAAGTATGCAGAGGGAGGGGTGGGGAG  
 24105 AGAGAACTGTGCAAGGAGTTTACCCAGGGTATGAAGAGGTAAAGAGGTCAGTATCAGGGAAGGAAGGTGCAAGAAAGGGTCAGGCTGGGAGG  
 24197 CTGGGCCACAGTTCAGTAAGATTACAAAGAAGGGCCTAGAACAATGAGGGCAGGCAGAAGGTGGCTGAAGGTGTAATTTTCATGGCAGGTTC  
 24289 TTTTCTAATCAGCTCCTCTAACCTCCTTCATCCTGTTGCCCCGGCTTTTGTTTTCCACTGTGACTAAGACATAGCCAAACAGGATATGACCG  
 24381 ACAGGAAGTTGTTTCAGTGCAAAAATAACTGATGTCTCATTCTGGAATATTATGGAAGGGCTCATTACTTACAGTGTGAGTGTGTAACCC  
 24473 AGGTTTTTCAGAATATTTTGTATAATCTTGGAGCTTATGTTTGTACATTTAGTACTGAACATCTGTATTGTTTTCTTATTAGAGAACACACTG  
 24565 TATTTACCTTAAACTGGTTCTTTTCTCCTATTGTCTATTATGGAACCAACAATTTTTATTGTAAATGTAACAGTGTGTAGCATCAGTCT  
 24657 TATAAATATTTTAGTTTGATACACAAACCGTAGTTCAAGTTAGTTAATTGATTTCTTCCCTAGAAAGTCAAGGAGTAACATAATCAGGTTAT  
 24749 AAACCTTCACTACTAGTTATTTAATAATTTATTTCTCTGGTTACATTTATATCTTAGGTGACATCAGAACATATATGTCACCTCCTTAAAGAT  
 24841 AGTGTGAAGAAAACCCACCTTATGTTTTCTTCCACAGCTTTTCTGTTTGTGAGCTTTTATTTTGTACTCAAAGAATAGCATCCAACTTTTA  
 24933 CTTTGGTTTCCCCATGTGGTTCTGAAAGAGAAGTAGAATTTCTTCTAAATCCGGAATTGCTCACATCCTTTTACCTTTTAACTTTGTTTTAA  
 25025 GCAAATGAACCTTATTGTTCCAGGTAAATCTTCCACAGTTGCATGCAGGGGAAAGTATGATGTCTCAGACTTTATAGTCTCATGGAGATGGAG  
 25117 TGAGGATCAAGGGCCATGCTCAGCAGAACTTGTGAGGACCCAGCAGTTTCACGGACACCTTTTCTTAATTTTTTAAACCAAGTCTATAATAAG  
 25209 TGCTTTCTTCCCTAGATTCCAATCCAGAAACAATATCATTGCATATTATACAAAGGAGCTGGCTAGGCTTGTGTCTGTGGGGTCAGCTGG  
 25301 TGTTCATTTCTGGGCTCCTTTGTGAAGAGGATGAAGTATGGTCTGAGAAGTTAGGTGTCTTGAAGTAGTGGAAATAAATCATGATAA  
 25393 CTCTTTAAATTAAAGATTATATATTTTGGCCTCAAACATTTTGCAAAGTCCTCCTATTCCAACCCAATCTGTTTAAATGACCCAACATTCA  
 25485 ACACATTGTTTCTGATAATTATCCTCAGAATAAGATGCTGTTGGCCATAATCTTTGTCTCTAGATTGTTTTATCTACTCGCAAATAAATTT  
 25577 AAGACACAGAGTATGCCTAAAGCCTACAGCAGACTTTCTGGAACTCTTGAATGTTTGGTCCATACTACTTCTTAAGACAAAGAAGAAAAC  
 25669 CTTGTCAGGGTGTGTCAATTAGTGCTTGAATGTAGGGTTTACAGGATGGGGTGGGGTGGGGGAATCGCCCTTGGTTTAGATGAATCATTCCTT  
 25761 TCCTTGTCTTCTCAGCAAAACACAGTTTCTACAGAGAACAGCTCTGCCATTGTGCATTTTCTGTCTCCATTTTCTCTCTCATTCCTCTCCA  
 25853 CGAAACCCAGAGTAGTCAGTGGGCTTTGGGCAGGAAAGTGGCAACAGGGTGTCTGGGGAAAAGCCAGTTGGCTCTTCTTACCATCACAAATAT  
 25945 AGACTGACCACAGGTTATTTTAAAGAGCAGAGCTGGTTTCCATCACTCTGAGAAGTGCTCAACTACAGACTTTGGGATGATATTTGTTATAGC  
 26037 TGTATTTTCTCCACTCTTAGATTGTGAAAGTACATATTACAAGTATTTATTTTATTATCTTTACTAAAATTTTAAATAAAAAGAAGCGTGCT  
 26129 TGCCGCAATAAGTAAAAATACCCAAAGTTGTTTAAAGAAAAGTTCACCTTTTCCCTTCATCCTCCATTCCCACATTCCTGAGAACACTGAAG  
 26221 TTAATAACCGGTTGCAATTCCCCCTTTCACCAAACTGATTGCTCATAGAAATATAGATAAACATATGTAAGGTTTTTAAGTTTTTTTAAATAAA  
 26313 AATATGTTTCATGATATATACATTATTCTGAAATTTTCTGTATCACTTAAAAATATTTCATAGATGTCCCTCTGGGCCAGTGGAAGATCTGGT  
 26405 TCCCCCTTACATACATATCAGCAAGCTGCATGATATTTCAACTATTGCTACTGCACAGTTTATTTCAGCCATCTCCCTATTAATGAACATTTA  
 26497 AGGTTTTTTTTTTCAGTTTTTAGCCTCTACAAACAGTACACAATAAACAACATGACATTAAATACTTGTGCTCTTATTTTCAGTAGGAGAAATT  
 26589 CCCCATGTGGAATTTTAAAGTCAAAGTTTATTGTGTTTTTAATGCTTTTAAACATTGCCAGGTTACCGTCCCAAAGGCTATAACAATTCAC  
 26681 ATTTCTGTTTCTCTGCATCTTACCAGACGAGTGTAAGTGGTATTTTACTGTGCTTTTCATTTATATTTTGTGCTGTTATTAGTGATATTTT  
 26773 CATATTTTCATATATTTATTTGCCATTTGTGTTTTTTCTTCTGACTTGCTTGTTCACATTGTTTACCTTGTTTTCTTCTGTCTTGTGTAGT  
 26865 GTAATAGTTTAGACTCTGAAGCCAGGCAACCTGAGTTAGAAGCCAGGCCTCTATTTTCATGATGTAGGTCTTTGGGCAAAGTACCTAACATTC  
 26957 ATGCCTTAGTGTTTTCTCTTTTAAATGAGCAGGGATAATAATAGTACCTGCCTCCTAAGGTTGTATAAAATTAAGTGGGCACTTAGGGTAAT  
 27049 ATCTAGCAGGTAGATATTGGCTATTATCAATAGTAGCTCTTATCGTTACTATTCTTCCAGATACTGTTTCTGACTCTGGGGCAAAGTCCTG  
 27141 CTACCCCTGAACCACATTTTTCTACCTCTTAGATTTTACTTGGTAATTCATCAGCCACTGTTGGGCATCCTCTGTGTTTAAATGCATCATCT  
 27233 TAGACCTTAGGAGGGATGGGAGGAACTTTAAAGAAGCCGAATTTGCTTTTTTATTTATCTTGTAGCAGAGCAATAGATGTATATTAGGTAGATT  
 27325 ACAAGCTTTTAGGTTATTTTTGCATCTAAAGCTGTCCCTTCTTTTCCAATAAATGATGTCTGTGGTAAAGAATATATCTGTTGGGTGTTAGT  
 27417 GACAAAATCAGAATGCTTTGTGTCTATTTTGGCTAGTAGTTAATTGTTTTCTTTTATTGTGTCTGCATTCCTATTTGTTCTTTAATTATAC  
 27509 CGAGCTCATTAGCAGTTATTCTTGCTTTATTCATTTCTTATCTCCTAGCATAGTCAGCTCAAGACAACAAGCATCTTTCAGAAAGCCACTAG

FIG. 12B  
6 of 10



27601 GAATTGCATCTACATTAAGAACCCACTCCTCTGCTCTAGCGTCTGAGAACATAACACAGTATTTGCCCTTTGTTGAAGGGCTCAAGCGGACGA  
 27693 TTCAGAAGTAGAATAGATAACTGTTGGTGGTGGCTTGCAGCACCTAGTACTGATGGTTTTGTTAGGAAACACCACAGGCAGTATACAATGTA  
 27785 AAGCAAGTTCTTTGGGTTTCAGTTAAATAGATTCCACTTGCACGTGTTCTCACTCTTTTGGTGTGAAAAATTAGGAAAGGTGATAGGCAGGAT  
 27877 AGAATAAAATGAAGTGGTCCCCTCTCCCATATGGAGAGCGCCTGCCTCCACCACAGACACATGTTTTGCCCTGGAAGCATAAACAGAAGATT  
 27969 GCAGGAACGCCCCCTTCACCTCCATAGCCTTCAGGCTCCCATCGATAGCATCAAGATAAACTTGGTGTGGCAACAGACTTGAGCCATCATCTT  
 28061 GTTTAACATTTTTTACCTGGAAGTGAAATGGAATCCAGAGAGGCTAAGTAGCTTGCACAGCTACTTAATTGAACTAGAACCAGAACCCAGTT  
 28153 TCTTAGTATTCTGGTGGCCATTTTATATAAACATAGACAGCTGATCATGGTAGTTTGGATCATTGCTAAAGACCTATTATATACAACAATCG  
 28245 TATGGCTAAGAAGATTAAGATGCTTATTTTCTCGTTATCCATGATTAGAAATGTAATAAGAGACTTTGTTCTGTCACCAACAGAAGAGACAA  
 28337 GACGCCTTTTCTGCTGTGGTTGTGAGTGCCCCATCAAGTGCTGAGATGTGACTCATGTCTTTGGGGAAAGTAGCAGATGAGAGAGTACATTC  
 28429 TATTCGACTCCATAGACTTAACTGTGGAGACTGAGAGTAGTGGAGAGCTCAGACTGACAAGAAGAACAGAAATATAGACTTAGAGGCACCAG  
 28521 GCATGAAAAATCATAAAGATGGAGATGACTTCTATCTTGTGAATGTCTGAAAAGCCTCTCTCTCTTTCTGTTCCCAAATCCTTCATCTCCAG  
 28613 CCCTCATCTCTGCCACCATCTGTGGTTTTCTTCCGTGGCCTCTCTTCTGTCCACCTCTTCAATGTTGGGCCCCCAGGATTTTGTCTTCAGA  
 28705 CTCACCTTCACCCCTCTCCCTCAGGACATCTCATCCACTCCCAAGGCTGGAATCCAGCCTTCCTCTCAGCCTTGCGCTTAAGTTTCATCCCT  
 28797 TTAGCTGTCTATTGGATGTTTCCAGTTGGAAATCCACAGGCATCTCTACCCAGGCGTCTCTACACAAAGATGATAAAAAATATTATCATCTT  
 28889 TTTGGCAGATTTGTTTTTCTGAGTTTTCTGTCTGTTCATGGATTACCATTCAGCAGGCTGCCCAAGCTAGAAATGTGGGATTTGTTAAT  
 28981 CGGCCCTGCTGTGAGACTGGGAGGCTGGTGTGATAACCCAAGAAAGACATTGTGGGCCTCTGGGTCTTATCCTTTTAAACCCATTGCCCTCAG  
 29073 CCCTGCCCGAGTGATGCTTCTGAAATGTGGAATCATTTTATTCTCTTACCTAAAAGGTTGCAGGATATTTTAGATTCTGAATGAAATCCCC  
 29165 AATCCTTTTTTTTTTTTTTTAAGTGAAAGCAAGTTTATTAAGAAAGTAAAGAAATAAAGAATGGCTATGCCATTGGCAAAGCAGCCCTGTGG  
 29257 GCTGCTGGTTGCCCATTTTTATGTTTTTCTTGATGATATGCTAAACAAGGGGTGGGAAATCCCAATCCTTGGCATTCAAATCCAGACT  
 29349 CATTTCTTTTCACTTTTTTTTTTTAATCATGCCCTGCCTTTCAGTTATGATGAGTGACTTGGTCATTGCTCAGGTGTGACTTGTCCCTTTGC  
 29441 ACCTGCTGTGCCTTCTCCTAGAAATGCGCTTTTCTCCTTTCTGGCCAAGTGTTCTTGTCTTCCAAATGGGCCTTCCCTTGGGAGGTGGTTT  
 29533 CTGACGACCACCCCTAGTCCAAGTCAGCTCCCACTGTACTTTAACTTTCTCTTGTCTTCCCTTATTACCTGTTGATATGCTCTCTCCCCAC  
 29625 CTGGTGTTCCTTGGGACTAGGGACTTCCTTCATTACATTTACATAACTTGAGGGCCTGGCTCATAAGAGGTGCTTAATGAATATTTATTG  
 29717 AATTAATTAGCATCTTGTCCTTCAAGATCAGCCATCATTTTCTCTATCTCATCATTCAAAATATATTCTTCTCTTCCCTTCTTGCACCC  
 29809 AGTCACAGACTGGACTCTATTAAATCCTGTCTATCATCTGGGCTCATTTCCATCCTCAGTGTCTGTCTGTGCATCCTTTTCATTACGCCAGG  
 29901 GATGTTACAGCTCGATTCTGCCCCCTTCAATCCAAGCCTGTCCCATATTCCATTACTTTATGAAGCCTTTCTTGACACACAGATGCTTAATTAT  
 29993 TCTCTTTTGTCTTCTTTGTGTTGACTTTGACTCTGCCACTGGTTGTGAGCTTCAGAAGGGCAGGGATCTCACCTTCACTTCTTTTCTCCT  
 30085 AGTGCTTTCTTTGTGTGCTGCACACTCCCTGGCACACACAGCGGCTCTCCAACACGAGGCAGAGCTTTCCAGCAGCCTCAACCTTCAGGACT  
 30177 GGGCAGCTTTTAAATGTATTTGGGCACCTTTGCAAGAAAAGGATTGTGTTAAATGTAATTTTCATGTATCTTGTATCAGAGTGCAAGAGAA  
 30269 GCAGAAATCAAATGAAAAGTGAAAGCAGAGTCAAATTTTTTATTTGGACCCAGATGCCAGGTAAGAACTATCTAAATGTTTAATATTTA  
 30361 AAACCAAATGTGGGAGAGAAAATCATCGATGGGCTTATTTGTTTATTTGTTTGCTTTGTTTATTTTGGAAAAACAAGCAAAATAACTATAGA  
 30453 AAATTTGGGGAAGAGGAAAAATAAAATGTATAATCTTATCACCATAGCATTACTATTGTGAATATTTGATATTCAATAGATGTTTGAAA  
 30545 ATTGGGAGAGATTTATTGAAAGACATTCTCAAGTTCACAAAGAACATCTAATTTACCTGTTAAATAAACCATCAGAAAACAACAGGTATCAC  
 30637 TGCAGTTGCCTGGGAGTCAGTGATAATTTCCGACTAGCCAGGCTCAGGCTCAAATACAAACCTTTTCCATTAACTCTAACGATAAGTACT  
 30729 TTTCTGTTTCTCACAACCTCATAACCATACGATGTGTGTTTATATGTCTATATTTTTTATTTGCTTTTAAAGAAGTTTTGTTTATCATT  
 30821 GTAAAAATATACATAATATAAAATTTACCATTTTAACCATTTTAAAGTGACGGTTCAGTGGCATTAATAACATTCTCATTGTTGTACAACCA  
 30913 TTACCACCATCCATTTCAGAACTTCTTCATTTTCCACACGGAACTTTGTATCAAATGATAACCTTCTTCTCTTCTTCCCCCATCCCCT  
 31005 AGTAACCTCTGTTCTACTCTGTGAACCTGCCTATTTTAGGAACCTCATAAATGTGGAATCATAACAGTATTTGTCTTTGTTCTGGCTTCTT  
 31097 AAACCTAACATGTTTTCAAGGTCAATCCATGTTGTAGCATGTGTGAGAATTTCTTCTCTTCTGTGGCTGAATATTCCATTGTATGTATATA  
 31189 CTACATTTTATATATCCTTGTAATCTGTTGATGGACACTTGGTTGGATACTTGATGGACATTGGTTTTGTTGTTTCATGATCATAATTTTCA  
 31281 AGCTCTGATTTTTTTCAGTTTCATCCATTGAGTAGGTATACCATCATGTCTTTTTTTTTTTTGTCTTTTTTTTTTTTTTTTGGAGGCAGAG  
 31373 TCTTGCTCTGTGCCCCAGGCTGGAGTGCAAGTGGTGAATCTCGGCTCACTGCAAGCTCCGCCTCCTGGGTTACACCATTCCTGCCTCAG  
 31465 CCTCAGCCTCCCGAGTAGCTGGGACCACAGGTGCCCACTACCACACCTGGCTAATTTTTTTGTATTTTTTTGTAGAGACGGGGTCTCACTGGG  
 31557 TTAGCCAGGATGGTCTCGATCTCCTGACCTGGTGAGCCGCCAGCCTCGGCCTCCCAAAGTGCTGGAATTACAGGCGTGAGCCACCGTGCCCCG  
 31649 GCCCATGTCTTTGACCATTTGTTATAAACTATGTGTGTAACACTATAAACCATAGAAACCGATTATATAATAGCAACACTATTGTGAGTAAA  
 31741 TAAGTGTATATAGCTTTTCCATATTTTATTCCGTTTCTTTGGATGCATTTATGATGTTTTTTTAAATAAGAGCATAACTTATTTATATGT  
 31833 TTACATTTTCTTTTAAAGAAGCTTGACTTCTCATCAGCTGAGCCAGAAGTGAAGTCATTTGAAGAGAAGTTTGGAAAAAGGATCCTTGICAA  
 31925 GTGCAATGATTTATCTTTCAATTTGCAATGCTGTGTTGCCGAAAATGAAGAAGGACCCACTACAAATGTAATTTTTTCAATTTAAAAATAAAC  
 32017 ATTAATAAAAAAATAGGCAGAGGTTTCAGATGTACCTTTACAGTGCAGCCTGGATAAGAAATCCTAGTCCCTGGTATCAAAGAGGTGCAGTG  
 32109 TTTGGATCAGGATATGGAGGTTGTTAGCCTGCAAGGACAGGATGTTCTGTGATGGAAGATGAGGGTGGCAGGTTTGTGCTCAGCTTTCCAGGA

FIG. 12B  
7 of 10





36801 ATACCTTTTAAATCGATATAGATAGGTGCATGGATGGGTCAATAGGCCTATTCTGTTTGTGTTTCAGAGACAAAGAGGATTTGAATGTGTAA  
 36893 AACTGAGAAATACATAAGCCCAGATTTTGA AAAAATCATTGGTAGAGTCACAGAGAGGATAGACACTGTCTGGAGAAGTGCTACCTGGAAC  
 36985 TGGCAGGGTGCACGGTAGTGTAGCTGCAGAGCTGGGATTCAAGGACCCAACCACATGCCTCCAGCTGGAAGTCAGGGCAATCCAGTGAGGC  
 37077 CTGGGGTGATCTTTATCTCTTGACTCTACTTGTTAAGCATTGTGACTTGTGTATATTGTTTCCCTAAGCACAGCCATTGGCTGGAAGTGT  
 37169 CTATGTAAATTGATTTAGTTGTCTCATCCCCATAGATGTTTCCATGTTTTTAGATAATGAGATTTCTGTTGGCTATAGCCAATGGAATAA  
 37261 TAATTAGACTTCTCATAGAACTAGACTTAAATAATGAATTGATTTTGGTGTGTTTGGAAACCCAGTTTAGAAATGTTGCTTTGCCATTTTCA  
 37353 GACATTGTTTAAGGATGCATCTGGAAATCTTGACAAAATGCCAGATTTTCTGCCATCTACAGGCAAGACAGCAATAAGCTATCCAATGATG  
 37445 ACATGCTCAAGTTACTTGACAGACTTTCGGAAGTGAGTTTCAAGGTCTTATTTCCACACCTGAAAAATAGAAGCTGTGTAGTGGGGAGGGAGG  
 37537 AACAGGGGAGCAGTCACCTAGGTGCTCGATTTAGACATCAGAGGGGATGGCAAATGAGCGTGAAGCATTTCCTCAAACCCCTTGAGAAGAAA  
 37629 GATGGGGTGAAAAATCAGAAGAATAACCAGTTAATTGAATTCTGTAGAGGATGTTTTGGGTGGTGTGTGAAGGGTGGACTGGGTAAGGATG  
 37721 AGCCTATGGTGGGGAGGAAACAGTTGAGGAACCTTGTCAAGAGGTGAGAAAGGACTCAGCAAAGCCACTGCAAGTGCAAACAGGAAGAAGGG  
 37813 GACAAATTCAGTCGTGCCAAAGAGATACGATGACTGGGTCTTGGCTTGGGGTGGTAATAGTCTAAGATAAAATACTTGCAAGGTTTCTAAC  
 37905 TTGGAAGTTTCTGGCACCAGTTGTTGCTTAGCTTGTGGCAGCATGCTCCACTCTGCCTCTGGCCTCACATGGTCCTCTTCTGTGTGTCTC  
 37997 TGTCTAAATTCCCTTCTTAGAACACTAGTGATAAAGCATCAGGGCCAAGCCTAGTGACCTCATCTTAAGTATGTCATCTGCAGAGACCCTG  
 38089 TTTCCAAATAGGTCACATTTATAGGTACAAAGGGTTGGGACTTCACCATGCTTTTGGAGGACACAATTCAACCCATAACAATGAGGCAAGA  
 38181 GGGAGCAAGGAATGTGTGCAACATCACAGGGCCGGCAGCTTCCCCAAGTCAGTCTCACCCGAGGGTCTGTGTTCTTAACCTCTATGCTGTTT  
 38273 TGCTGCTACATCCTAAAGAGTTCACTCTGAACCTTTGAAACTGATTTTCTTTGCTAGGGAGATGGGTCTTAGAATTTTTCTGGGGAAATTCT  
 38365 GGGAAATGTGAAAGAGCTGAGGGCGCTAGAAGATGTGAAGTGAAAAGAATAGCTGAGAGCCAAATGCTAACTATTCATGCCAAAGGTATCCT  
 38457 TGTTTTTTTTTTTTTTTGTGCATATCAAATAGCAATCTTATCAGTTTGTCTAGAACTCAAGAATGATTGCTTAGCTTTCTTTAACCTTATT  
 38549 TTACCTTTTTCTTATCTGTCTTCAGTAGTAGGAATAGAAACGATATGAGTCATAGAAACAGGCTCAATAAGTTCTGAAAACACAGAGACGTG  
 38641 TTCCTAATCAGAATCCAATCACGTCCATGTCAGCAGGCGGCTTCAGCCTTCACAGCGACGTGAAATCCCTTGTCAAGAGGCTCAAAAAGGTA  
 38733 GAAAGGATTCTCAAGGTCTCTTTCAGTTATGTGATTATACAGTTTGTGACTGTCTTGATGTTTCCCTGTTTGGAGCTTTAATGAGAAGTGC  
 38825 AACCTCAGTTTTGCTAACATGCAGCTAAGGTTGGCCTGTTTCAGCAAAGCAGTGTGCATGCCCGCTGGGCTGATTGGAATGAACCTTTCACA  
 38917 GCTCACGTAGGGAATTGGAGAAGGGGGAGAGGAGGATACTGGTGAAGGATGAGGCCTGCTGGGTAGCCTTCCAGGGTTCTTGACCATATA  
 39009 GGTGCCCCAAATTCCCAGTCACTATCTGACAGTTTTATGACCTGGTAAGGACACAGGTCCTGGCCAGGGAGTGCCCTGGATCCCTATGAAT  
 39101 CTGTTATTTCATGAAAGACTAAATAAAAAGAATAGTACCCTATTTTTACTTTTAAATCATAGAGGTTCTTTAGTTTACAAACATAATACATGTT  
 39193 CATTTTAGAAATTTTGAGAAATACAGAAGAATAAAAGGATGAAAAAGGTTTACTACTAGTTTAACTTTCGTGGTGAACTTTGGAGAAGT  
 39285 TTTTTTTTTTTTTTTGAGATGGAGTCTCACTCTGTCCCTCAGGCTGGCGTACAGTAGCAGGATTTAGCTCACTGCAACCTCCGCCTCCCGAG  
 39377 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTATAGGCGCCACCACCTACGCCTGGCTAATTTTGTATTTTAGTAGAGA  
 39469 TGGGGTTTCACCATATTGGCCAGGCTGATCTCAAACCTGCTGACCTTGTGATCTGCCCCGCTCAGCCTCCCAAAGTGCTGGGATTACAGGCAT  
 39561 GAGCCACCGTGCCAGCCTAGGGGGGAACATTTTTTTTTTACGTTTTATTCCTTTACATTTTATTTTAGTTTATCTTATGTAGCTATGATCAT  
 39653 ACTAAATATGTAATATTTCCCTGCACAACCTCAAGTATTTTCTGAAAGTGTTATATATACATTTTATAGACATCATTTTAAATGCATAAATA  
 39745 TTATAATAGTCCATTGAGATAGACCATAGATTATTTAACTCTTCCCCCATTTTTTGGACTTTTTTTTTTTTTTTTCCGAGATGGAGTCTCGCTCT  
 39837 GTCGCCAGGCTGGAGTGCAGTGGCACCATCTCAGCTCACTGCAACCTCCGCCTCCAGGTTCAAGCAATTCTCCTGCCTCAGCCTCCTGAGT  
 39929 AGCTGGGACTCCTGAGTAGCTGAGTAGCGCATGCTGCCACGCCCGCTAATTTTTTTTTTTTTTTTTTTTTTTTGTATTTTAGTAGAGA  
 40021 TACTAAATATCTCACCATCTTGCCAGGCTGGTCTCAAACCTCCTGACCTCAGGCAATCTGCCCCGCTCAGCTTCCCAAAGTGCTGGGATTAC  
 40113 AGGTGTGAGCCACCATGCCAGCCATTTTTTGACTTTTAAATGTGTTTCTGATTTTTCAGAAATATACCTATAAGCCACAGTTAGAATCTTTA  
 40205 AAAAAATCTTCTCTATTGGTAGTGGGTAATATATTATCATACATACTATATTATCATATAGTAATTATTGTCATTTTTTGGAGTTTCAAGAAA  
 40297 ATTTCAATCTTACTAATTTTTTCAAAAACAGTCACCTTTAGTTGGATAGATTTCAATATTTTCTTCCGCTCAACTACCATGCAACTCTTAA  
 40389 TAACCATGAGGTGGGTCTGCGTGTACTTAGGAAAGTGAATACTATATTATTAAGGAAGAAAAATATATCTGTATTACTATATTTTTTGA  
 40481 AAGAAAATATATATTTCTTTTGTATGTAATGAAGAAATGGATAAGCAAGTAGCTATCTAGATGGAAAGATAGGCATAAAATAGCTATTTA  
 40573 GGATATATGCCAAATAATCATGGTTATCTCTGAGGGATGGGTTGATGGGTGATATTTACCTTCCACTTTATAACATTCTGTCAATTTTATA  
 40665 TGAACTTTAAAAAACTAACACTTTTATATTCAGACAAAACAAACATGAAGTTTTTATATGTGATGGAGGTTGGAGCCCTGTCTCAGAAG  
 40757 TTACTTCTTAGGCTGGTTAGCTTGAGACTTCCCCACAGTGGCGGCTCCTCAGGGGCAGCCCCAGTGCATGGTCTGTCTTCAGTGGAGGCTG  
 40849 GGGAGTGGGGCTTCACATGGTCACTAATTTGAAAGTGATGGGAGCAGAAAGCCTGTGGCCAGGCAGAAAGGAGCCAGGGAAAACCAAGTGT  
 40941 GAGTTCTCTTCTGCACACCCTTCTTCATGCATGTGCTCAGCAGGAGGGCATTGGTGTGAAGGGTGTGCTCCAGGTGGCCAGTTAGAGACCC  
 41033 AGAAACCTGAAAACAGGGATCCGATGGTGACAGCATAGAAGACACAGCAGGATAAGTGAGGCCACGCTCCTCAATAAGTATTCAAAGAACT  
 41125 TTGGTGCCCACTCCCCGTATTCTTCACAACAGAGTTAGGGGACGTGGAGGATTCCTTTTTTCATTTTTTAAAAATCTTGCATTGCTATTTTT  
 41217 CTTTCTCTGTATATTTTACAGGAAATAATCTCATGTGCTGAGTGGCCTGGGCACCGGCTTGGATCCAAAGCTATTGTTTCTACCCCATGATTG  
 41309 TCTCAAATGTTATTTAATAATGCATGAAAAAATTTCTTCACGCTGTCTCAGTCTTAACAAAACAGCTGCCAAAGCTCATAAGCCACTTTC

FIG. 12B  
9 of 10

41401 CTTTTCCCTTGCAATAATTACCCAGGGATATGTTCCAAGATTTAGTAAGAAAGCGATTCTGTCCGATAGATGATATTGCTAACATTTTATA  
 41493 AGAAGAGAGACTTGGTACTTTGTATTTGATTTGTTTCATGGTGGTATCTCATGGATAAGATGGTATCTCATCTTTTCCAACCTTCTGCAGGAAA  
 41585 TGCGAAGACATGAAGGCAAAGTATAAAAATAGAACGTTTCTTTAAAACGTAGACCTTTTAAATGGTACTACGTTGGATAGTTTAGGTAATA  
 41677 ATACTACTAAAGTTTTTGCCTATGCAGCTTAATGTGTCTGTGTTTATTTGTACACTCATCTTCTTTGCATCCAGGTTTACAGTCTTACCCC  
 41769 GATTTGCTCTGGTTACACTGCACTCAAGCCAAGTAGGGCTGCTTGACTTTCTCTAAACCCACTGGGGACTTCCCTCTGCCATGCTTTTCTC  
 41861 TCTGCCCCAAATTGTGTCCCCCTTCCCTGCCTCATCAAGCAGCACATAAATCACAAACACATGCAGCATAACACTTCCCCTTTTCTTTGTCTT  
 41953 TCTCAGGGAACCTTACTCATCTTTCAAAGCCAGTCTGTGGCTCACTTCTGTGCTGGGAGTCTTGGAGGCGGTACTTGGCTTCTCTGCCTG  
 42045 AGCCGCCTCCTCTTTTTTAAGGGTGGATAATAACAGCCCCCTGCCCCCTAAAACCGTGGTGGGGAATAAATGCAAAGGCATTAAGGTGATTTC  
 42137 TCCCACCATGAATACTGATCTCATCCCGTGTTCCTCTCGATAGATCTAGATACTCTGCCTTCTGGTAGAGGTTTGTACATACTCTGTGAAA  
 42229 GTGATTGCCCTCATATGCCGTAAGTAGCTTACAGTGTCTACTGGACTTTTGGCTTCTTGAGGAAAGAAATTATGTCTTGTGTTGCATTCTCCTCC  
 42321 ATGGTCCTGAGTACATACATTGCAGCATATCCTAAGCACTTGATAAATGCTTATTGAATTTTCTTCTTAGACATAAACTCAGTGGTTTTTGT  
 42413 TGAAACAAAAATATCTCAAATTTCTTTCAATCATATATAGTTGTTTTTTTTTAAGTGACACCAAAGCTTTTAGGGAATATTTCCTTTTACAA  
 42505 AACACAGTTAGAAGATTAACTCACCACCAATAGCAGTCCAACATACCTGTATTGCCAGCTAATCATTTTAACGAGCCAATACAGGAAGTC  
 42597 AGGAAGGGAAGACCGGCTGCAGAAACACTTAGATAAGGACCCCAAATCTGTGTCATGGGAGGACTGCTAGTTGATGATACCATTCCCATT  
 42689 CCTCTGTGGGAATTGTTGAGTCAGCAGAAATGGATGGGCAGTGGGAAGGGAATTTTCCCTAAGAGAGAGTTTGAGCCTCACTTCTACATTC  
 42781 ACACAGAGACAGGAGCAGTTCCAGAGGCCAGGCATCCTGCAAGTGTCTGTATTGCATGCTTACTTAATTCGTGTAATTTTAAGATGAGTTT  
 42873 TCATGTTCAAGGATTATTTTATAAATTTTGCATAGAATATAGGTACTCTTTAGCAAAACAAAGCAAAAAAACCAAACTATTCTCAGTCATG  
 42965 AAAGAATTCAGTTTGTGTAACACGCACACAACCACCACTTTGGAAGTGCATAAAAAGGCAGTAAATCTTTATTGCCTGTGAGTGTGTTGATG  
 43057 TCTAATAAACCAGATTCAACATAAACCATAAACTTTTGAAATGGGTTTGAGATTGGGTTTTTAAAACTTAAAGCTGGCAAAAAAAAACA  
 43149 ACTTTTAAAAGCCCATGTGCTACATAATATGGAATAAATCAGAAATGTGCTTGGAAACACATGGAAAGAACGTCTTTACAGAAGCAGCAA  
 43241 CTAGAAGTAAATCTCTCAGCAGAGGGAGGAAATAGAATAAGAAATAACTATAGTTAGGCACAGAAGGACACAATACTATAGGAAGATTT  
 43333 CCAGTGAAGATCATTTAATTAATAATATGTTGCTTAGAAACGTATTTAATTGTGTTCCACCTCTCTCAAAAATTTATATGTGGAGGATGTTG  
 43425 GAGTGATCTTAAAAATGGTGATGAAGATGCCCTGTTTCATTATAGGTGGAAATAATTAGGAGGGGGTGAAATCCATTACCCTTGCATACTTAC  
 43517 TTATATTTAAAAGTATAATTTGTAATAAA

hCLASP4	-----MFPMEDISISVIGRQRRTVQ-----	20
hCLASP5	-----MTHLNSLDVQLAQELG-----	16
hCLASP3	-----MAERRAFAQKISRTVAAEVRKQISGQYSGSPQLLKNLNIVG	41
hCLASP2	-----MLLFPYDDFQTALRRQGRYICS-----	23
hCLASP7	-----MAASERRAFAHKINRTVAAEVRKQVSRERSGSPHSSRRRCSSSL	43
hCLASP1	MSFRGKVKREPSEFWKKRRTVRRVIQEEFHRFSSQEKPRLLEPLDYETVIEELEKTYRN	60
... :		
hCLASP4	-----STVPEDA EKRAQSLFVKECIKTYSTDWHV VNYK	53
hCLASP5	-----DFT	19
hCLASP3	N-----ISHHTTVPLTEAVDPVDLEDYLITHPLAVDSGPLRDLIEFP	83
hCLASP2	-----TVPAKAE EEAQSLFVTECIKTYNSDWHLVNYK	55
hCLASP7	G-----VPLTEVVEPLDFEDVLLSRPPDAEPGPLRDLVEFP	79
hCLASP1	DPLQDLLFFPSDDFSAATVSWDIRTLYSTVPEDA EHK AENLLVKEACKFYSSQWHV VNYK	120
::		
hCLASP4	YEDFSGDFRMLPCKSLRPEKIPNHVFEIDEDCEKDED-----SSSLCSQKGGVIKQG	105
hCLASP5	DDDL DVVFTPKCRTLOP-SLPEEGVELDPHVR-----DCVQTYIREWLI	63
hCLASP3	PDDIEVVYSPRDCRTLVS-AVPEE-SEMDPHVR-----DCIRSYTEDWAI	126
hCLASP2	YEDYSGEFRQLPNKVVKLDKLPVHVYEVDEEVDKDED-----AASLGSQKGGITKHG	107
hCLASP7	ADDLELLLQPRECRTEP-GIPKD-EKLDAQVR-----AAVEMYIEDWVI	122
hCLASP1	YEQYSGDIRQLPRAEYKPEKLPSHSFEIDHEDADKDEDTTSHSSSKGGGGAGGTGVFKSG	180
:: . : * . : * .		
hCLASP4	WLHKANVNSTIT--VTMKVFKRRYFYLTQLPDGSYILNSYKDEKNSKESK-GCIYLDACI	162
hCLASP5	VNRKNQGSPEIC--GFKKTGSRKDFHKT-LPKQTFESETLECSEPAQA--GPRHLNVLC	118
hCLASP3	VIRKYHKLGTGF--NPNTLDKQKERQKG-LPKQVFESDEAPDGN SYQDDQDDLKRRSMSI	183
hCLASP2	WLYKGNMNSAIS--VTMRSFKRRFFHLIQLGDGSYNLNFYKDEKISKEPK-GSIFLDSCM	164
hCLASP7	VHRRYQYLSAAY--SPVTTDTQERERQKG-LPRQVF EQDASGDERSGPEDSND SRRGSGSP	179
hCLASP1	WLYKGNFNSTVNNTVTVRSFKKRYFQLTQLPDNSYIMNFYKDEKISKEPK-GCIFLD SCT	239
: : : : * : : . : .		
hCLASP4	DVVQCPKMRRHAFELKMLDKYSHYLA AETE QEME EW LITLKKI IQINTDSL VQEKKETVE	222
hCLASP5	DVSGKGPVTACDFDLRSLQPKRL ENLLQQVSAEDFEKQNEEARRTN-----RQAE	169
hCLASP3	DDTPRGSWACSI FDLKNSLPDALLPNLLDRTPNEEIDRQND DQRKSN-----RHKE	234
hCLASP2	GVVQNNKVRRFAFELKMQDKSSYLLAADSEVEME EWITILNKILQLN-----FEAAMQEK	219
hCLASP7	EDTPRSSGASSI FDLRNLAADSL LPSLLERAAPEDVDRRNETLRRQH-----RPPA	230
hCLASP1	GVVQNNRLRKYAFELKMNDLTYFVLAAETESDMDEWIHTLNRI LQISPEGPLQGRRSTEL	299
* : * : : : : :		
hCLASP4	TAQDDDETSS----QGKAENIMASLERSMHP ELMKYGRETEQLNKLSRGDGRQNLFSFDSE	278
hCLASP5	LFALYPSVD----EEDAVEIRPVPECPKEHLG-----N-----RILVKLLTLKFEIE	212
hCLASP3	LFALHPSPD----EEEP IERLSVPDIPKEHFG-----QRLLVKCLSLKFEIE	277
hCLASP2	RNGDSHEDD----EQSKLEGSGSGLD SYLPELAKSAREAEIK---LKSES RVKLFYLDPD	272
hCLASP7	LLTLYPAPD----EDEAVERCSRPEPPREHFG-----QRILVKCLSLKFEIE	273
hCLASP1	TDLGLDSL DNSVTCECTPEETDSS ENNLHADFAKYLTETEDTVKTTRNMERLNLFSLDPD	359
: : : : : :		
hCLASP4	VQRLDFS----GIEPDIKP-FEEKCNKRFLVNCHDLTFN ILGQIGDNAKG PPTNVEPFFI	333
hCLASP5	IEPLFAS----IALYDVKERKKISENFHCDLNSDQFKGFLRAHTPSVAASSQARS AVFSV	268
hCLASP3	IEPIFAS----LALYDVKEKKKISENFYFDLNS EQMKGLLRPHVPPAAITTLARSAIFSI	333
hCLASP2	AQKLDFS----SAEPEVKS-FEEKFGKRILVKCNDLSFNLQCCVAENE EGPTTNVEPFFV	327
hCLASP7	IEPIFGI----LALYDVREKKKISENFYFDLNSDSMKGLLRAGH THPAISTLARSAIFSV	329
hCLASP1	IDTLKLQKKDLLEPESVIKPFEEKA AKRIMIICKALNSNLQGCVTENENDPITNIEPFFV	419
: : : : : : : : * :		

FIG. 13  
1 of 6

hCLASP4	NLALFDVKNCKISADFHVDLNPPSVREMLWGSSTQLASDGSP---KGSSPESYIHGIAE	390
hCLASP5	TYPSSDIYLVVKIEKVLQOQD----IGDCAEPYTVIKESDG-----GKSKE-KIEKLKL	317
hCLASP3	TYPQDVFLVIKLEKVLQOQD----IGECAEPMIFKEADA-----TKNKE-KLEKLKS	382
hCLASP2	TLSLFDIKYNRKISADFHVDLNHFVSRQMLATTSPALMNGS-----GQSPSVLKGILHE	381
hCLASP7	TYPSPDIFLVIKLEKVLQOQD----ISECCEPYMVLKEVDT-----AKNKE-KLEKLRL	378
hCLASP1	SVALYDLRDSRKISADFHVDLNHAAVRQMLLGASVALENGNIDTITPRQSEEPHIKGLPE	479
	. . *: *: . : . : . . . :	
hCLASP4	SOLRYIQQGIFSVTNPHPEIFLVARIEKVLQGNITHCAEPIKNSDPVKTAQKVHRTAKQ	450
hCLASP5	QAESFCQR-----LGKYRMPFAWAPISLSSFFNVSTLEREVDVDVSVVGRSPVGERRTLA	372
hCLASP3	QADQFCQR-----LGKYRMPFAWTAIHLMNIVSSAGSLERDSTEVEISTGERKGSWSERR	437
hCLASP2	AAMQYPKQGI FSVTCPHPDIFLVARIEKVLQGSITHCAEPMKSSDSSKVAQKVLKNAKQ	441
hCLASP7	AAEQFCTR-----LGRYRMPFAWTAHLANIVSSAGQLDRDSD----SEGERRPAWTDRR	429
hCLASP1	EWLKFQKQAVFSVSNPHSEIVLVAKIEKVLGMNIASGAEPYIKNPDSNKYAQKILKSNRQ	539
	: : : . : : . :	
hCLASP4	VCSRLGQYRMPFAWAARPIFKDTQGSLLDGRFSPLYKQDSSKLSSDILKLLSEYKKPE	510
hCLASP5	QSRRLSERALSLEENGVS NFKTS-----TLSSVSSFFKQEGDRLSDEDLFKFLADYKRSS	427
hCLASP3	NSSIVGRRSLERTTSGDDACNLTSFR-PATLTVTNFFKQEGDRLSDEDLYKFLADMRRPS	496
hCLASP2	ACQRLGQYRMPFAWAARTLFKDASGNLDKNARFSAIYRQDSNKLSNDDMLKLLADFRKPE	501
hCLASP7	---RRGPQ--DRASSGDDACSFSGFR-PATLTVTNFFKQEAERLSDEDLFKFLADMRRPS	483
hCLASP1	FCSKLGKYRRAFAWAVRSVFKDNQGNVDRDSRFSPLFRQESSKISTEDLVKLVS DYRRAD	599
	. : : : * : : : * : * : : : : . . .	
hCLASP4	--KTKLQIIPGQLNITVECVVDLSNCITSSYVPLKPFE-KNCQNITVEVEEFVPEMTKY	567
hCLASP5	SLQRRVKSIPGLLRLEISTAPEIINCCLTPEMLPVKPPF-ENRTRPHKEILEFP--TREV	484
hCLASP3	SVLRRLRPITAQLKIDISPAPENPHYCLTPELLQVKLYP-DSRVRPTREILEFP--ARDV	553
hCLASP2	K-MAKLPVILGNLDITIDNVSSDFPNYVNSSYIPTKQFETCSKTPITFEVEEFVPCIPKH	560
hCLASP7	SLLRRLRPVTAQLKIDISPAPENPHFCLSPELLHIKPYP-DPRGRPTKEILEFP--AREV	540
hCLASP1	R-ISKMQTIPGSLDIAVDNVPLEHPNCVTSSFIPVKPFNMMAQTEPTVEVEEFVYDSTKY	658
	: : : . * : : . . : . . : * : : * : * :	
hCLASP4	CYPFTIYKNHLYVYPLQLKYDSQKTFKARNIAVCVEFRDSDSDASALKCIYGKPAGSV	627
hCLASP5	YVPHTVYRNLLYVYPQRLNFVN--KLASARNITIKIQFMCG-EDASNAMPVIFGKSSGPE	541
hCLASP3	YVPNTTYRNLLYIYPQSLNFAN--RQGSARNITVKVQFMYG-EDPSNAMPVIFGKSSCSE	610
hCLASP2	TQPYTYITNHLYVYPKYLKYDSQKSFARKNIAICIEFKDSDEEDSQPLKCIYGRPGGPV	620
hCLASP7	YAPHTSYRNLLYVYPHSLNFSS--RQGSVRNLAVRVQYMTG-EDPSQALPVI FGKSSCSE	597
hCLASP1	CRPYRVYKNQIYIYPKHLKYDSQKCFNKARNITVCIEFKNSDEESAKPLKCIYGKPEGPL	718
	* * * : : * * : : . . * : : : : : . * . : : : * : : .	
hCLASP4	FTTNAYAVVSHHNQNPEFYDEIKIELPIHLHQKHLLFTFYHVSCBINTKGTTKKQDTVE	687
hCLASP5	FLQBVYTAVTYHNKSPDFYEEVKIKLPKALTNNHLLFTFYHISCOQ-----KQGASVE	595
hCLASP3	FSKRAYTAVVYHNRSDFHEEIKVKLPATLTDHHLLFTFYHVSCQ-----KQNTLE	664
hCLASP2	FTRSAFAAVLHHHNQNPEFYDEIKIELPTQLHEKHLLLTFFHVSCDNSSKGSTKKRDVVE	680
hCLASP7	FTRBAFTPVVYHNKSPDFYEEFKLHLPACVTENHLLFTFYHVSCOP-----RPGTALE	651
hCLASP1	FTSAAYTAVLHHSQNPDFSDEVKIELPTQLHEKHLLFSFYHVTCINAKANAKKKEALE	778
	* : : * : * : : * : * : * : : : : : : : : : : : : : : : : *	
hCLASP4	TPVGFAWVPLLKDGRITFEQQLPVSANLPPGYLNLNDAESRRQCNVDIKWVDGAKPLLK	747
hCLASP5	TLLGYSWLPILLNERLQTGSYCLPVALEKLPPNYSMHSAEKVPLQNPPIKWAEGHKG VFN	655
hCLASP3	TPVGYTWIPMLQNGRLKTGQFCLPVSLKPPQAYSVLSPEVP---LPGMKWVDNHKG VFN	721
hCLASP2	TQVGYSWLPLLKDGRVVTSEQHIPVSANLPSGHLGYQELGMGRHYGPEIKWVDGGKPLLK	740
hCLASP7	TPVGFTWIPLLQHGRRLRTGPFCLPVSDQPPPSYSVLTDPDA---LPGMRWVDGHKG VFS	708
hCLASP1	TSVGYAWLPLMKHDQIASQEYNIPIATSLPPNYLSFQDSASGKHGGS DIKWVDGGKPLFK	838
	* : : : * : : . : *	

FIG. 13  
2 of 6



hCLASP4	FKSHLESTIYTQDLHVHKFFHHCQLIQS-----GSKEVPGELIKYLKCLHAM	794
hCLASP5	IEVQAVSSVHTQDNHLEKFFTLCHSLESQVTFPIRVLDQKISEMALEHELKLSIICLNSS	715
hCLASP3	VEVVAVSSIHTQDPYLDKFFALVNALDEH-LFPVRIGDMRIMENNLLENELKSSISALNSS	780
hCLASP2	ISTHLVSTVYTQDQHLHNFFQYCQKTES-----GAQALGNELVKYLKSLHAM	787
hCLASP7	VELTAVSSVHPQDPYLDKFFTLVHVLEEG-AFPFRLKDTVLEGNVEQELRASLAALRLA	767
hCLASP1	VSTFVVSTVNTQDPHVNAFFQECQKREKD-----MSQSPTS NFIR SCKNLLNVE	887
..	*:: ** :.. ** : ..	:
hCLASP4	EIQVMIQFLPVILMQLER-----VLTNMTH-----EDDVP	824
hCLASP5	RLEPLVLFLHLVLDKLFQLSVQPMVIAGQTANFSQFAFESVVAIANSLHNSKDLSDQHG	775
hCLASP3	QLEPVVRFLHLLLDKLI LLVIRPPVIAGQIVNLGQASFEAMASI INRLHKNLEGNHDQHG	840
hCLASP2	EGHVMIAFLPTILNQLER-----VLT-RAT-----QEEVA	816
hCLASP7	SPEPLVAFSHHVLDKLVRLVIRPPIISGQIVNLGRGAFEAMAHVVSIVHRSLEAAQDARG	827
hCLASP1	KIHAIMSFLPIILNQLER-----VLVQNE-----EDEIT	916
.	:: * : * : *	:
hCLASP4	INCTMV-LLHIVSKCHEEGLDS-----YLRSEFIKYS-----FRPEKP	860
hCLASP5	RNCLLASVHYVFRLEPEVQRDVPKSGAPTALLDPRSHTYGRTSAAAVSSKLLQARVMSS	835
hCLASP3	RNSLLASYIHYVFRLEPNTYPNSSSPG-PGGLGGSVHYATMARSAVRPASLNLNRSRSLN	899
hCLASP2	VNVTRV- I IHVVAQCHEEGLES-----HLRSYVKYA-----YKAEPY	852
hCLASP7	HCPQLAAYVHYAFRLPGTEPSLPDGAPP----VTVQAATLARGSGRPASLYLARSKSISS	883
hCLASP1	TTVTRV-LPDIVAKCHEEQLDH-----SVQSYIKFV-----FKTRAC	952
.	.	:
hCLASP4	SAPQAQLIH-----ETLATTMIAILKQS-----	883
hCLASP5	SNPDLAGTHSAADEEVKNIMSSKIADRNC SRMSYYCSGSSDAPSSPA-----	882
hCLASP3	SNPDISGTPTSPDDEVRSIIGSKGLDRSNSWVNTGGPKAAPWGSNPSPSAESTQAMDRSC	959
hCLASP2	VASEYKTVH-----EELTKSMTTILKPS-----	875
hCLASP7	SNPDLAVAPGSVDDEVSRILASKLLHEELA-LQ-----	915
hCLASP1	KE---RPVH-----EDLAKNVTGLLKS-----	972
.	.	:
hCLASP4	-----ADFLSINKLLKYS-----WFFFEIIAKSM	907
hCLASP5	-----APRPASKKHFEELALQ-----MVVSTGMVKSM	910
hCLASP3	NRMSSHTETSSFLQTLTGRLPTKKLFHEELALQWVVC SGSVRESALQQAWFFFELMVKSM	1019
hCLASP2	-----ADFLTSNKLLRYS-----WFFFDVLIKSM	899
hCLASP7	-----WVVS SAVREAILQHA-----WFFFQLMVKSM	942
hCLASP1	-----DSPTVKHVLKHS-----WFFFAILIKSM	995
.	.	:
Cadherin Cleavage		
hCLASP4	ATYLLEENKIKLRFGRFPETYHHVLHSLLLAIIPHVTIRYAEIPDE---SRNVNYSLAS	964
hCLASP5	AQHVNMDKRDSEFRRTRFSDRFMDITTIVNVVTSEIAALLVKPQKENEQAEMNISLAF	970
hCLASP3	VHHLVFNDKLEAFRKS RFPERFMD DIAALVSTIASDIVSRFQKDETEM---VERLNTSLAF	1076
hCLASP2	AQH LIENSKVKLIRNQRF PASYHHAETVVNMLMPHITQKFGDNPEA---SKNANHSLAV	956
hCLASP7	ALHLLLGQRLDTERKLRFPGRFLDDIT ALVGSVGLVITRVHKDVEL---AEHLNASLAF	999
hCLASP1	AQH LIDTNKIQLERPQRF PESYQNELDNLMVLS DHVIWKYKDALEE---TRRATHSVAR	1052
.	:: : . * *. : . : : . . . . *	:
hCLASP4	FLKRCLTLMDRGFIENLINDYISGFSPKDP-----KVLAEYKFEFLQTCNHEHYIPLNL	1019
hCLASP5	FLYDLLSLMDRGFVENLIRHYCSQLSAKLSNL---ETLISMRLFLRILCSHEHYLNLNL	1027
hCLASP3	FLNDLLSVMDRGFVFSLIKSCYKQVSSKLYSLPNPSVLVSLRLDFLRIICSHEHYVTNL	1136
hCLASP2	FIKRCFTFMDRGFVFKQINNYISCFAPGDP-----KTLFEYKFEFLRVVCNHEHYIPLNL	1011
hCLASP7	FLS DLLSLVDRGFVFSLVRAHYKQVATRLQSSPNPAALLTLRMEFTRILCSHEHYVTNL	1059
hCLASP1	FLKRCFTFMDRGCVFKMVNNYISMFS S GDL-----KTL CQYKFD FLQEV CQHEHF I PLCL	1107
*	:::*** :* . . . . *	:

FIG. 13  
3 of 6



Cadherin EC motif

hCLASP4	PMAFAKPKLQR-----VQDS--NLEYSLSDEYCKHHFLVGI	LLRETSI	1060
hCLASP5	FFMNADTAPTSP--CPSISSQNSSSCSSSFQDQKIASMFDLTSEYRQQHFLTGI	LFTELA	1085
hCLASP3	PCSLTTPPASPSVSSATSQSSGFSTNVQDQKIANMFELSVPFROQHLAGI	VLTELA	1196
hCLASP2	PMPFGKGRIQR-----YQDL--QLDYSLTDEFCRNHFLVGI	LLREVGT	1052
hCLASP7	PCCPLSPPASPSVSSSTTSQSSTFSSQAPDPKVTSMFELSGPFROQHFLAGI	LLTELA	1119
hCLASP1	PIRSANIPDPLTP-----SES----TQELHASDMPEYSVTNEFCRKHFLIGI	LLREVGF	1157
		.....	
hCLASP4	ALQDN----YEIRYTAISVIKNLLIKHAFDTRYQHKNQQAKIAQLYLPFVGLLEN	IQRL	1116
hCLASP5	ALDAEGEGISKVQRKAVSAIHSLLSSHDLDPKCVKPEVKVKIAALYLPVGI	ILDALP--	1143
hCLASP3	ILDPDAEGLFGLHKKVINMVHNLSSHDSDPYSDPQIKARVAMLYLPLIGI	IMETVP--	1254
hCLASP2	ALQEFR----EVRLIAISVLKNLLIKHSFDDRYASRSHQARIATLYLPLFGLLI	ENVORI	1108
hCLASP7	ALEPEAEGAFILHKKAISAVHSLLCGHDTDPYAEATVKARVAELYLPLLSI	ARDTLP--	1177
hCLASP1	ALQEDQ----DVRHLALAVLKNLMAKHSFDDRYREPRKQAQIASLYMPLYGML	LDNMPRI	1213
	..	.....	
hCLASP4	AGRDTLYSCA-----AMPN-S----ASRDEFPCGFTSPANRGS	LSTDKDTAYGS	1160
hCLASP5	-----QL-----CDFTVADTRRYRTSGSD-----		1162
hCLASP3	-----QLY-----DFTETHNQGRPICIA	TDD--	1276
hCLASP2	NVRDVSPFPVNAGMTVKDES	LALPA-VNPLVTPQKGSTLDNSLHKDLLGAISGIASPYTT	1167
hCLASP7	-----RLH-----DFAEGPGQRSRLASMLDS	TE	1201
hCLASP1	YLKDLYPFTVNTSNQGSRDD	LSTNGGFQSQTAIKHANSVDTSFSKDVLSIAAFSSIAIS	1273
		..	
hCLASP4	FQ-NGHGIKREDSRGS	LIPEGATGFPDQGNTGEN-----TRQSSTRSSVSQYNRLDQYE	1213
hCLASP5	-----EEQEGAGAINQNVALAIAGNNFN	LKT-----SGIVLSSLPYKQYNMLNADT	1208
hCLASP3	-----YESESGSMISQTVAMAIAGTSVP	QLTR----PGSFLLTSTSGRQHTTFSAES	1324
hCLASP2	STPNINSVRNADSRGS	LISTDSGNSLPERNSEKSNSLDKHQQSSTLGNSVVRCDKLDQSE	1227
hCLASP7	-----GEGDIAGTINPSVAMAIAGG	PLAPGSR---ASISQGPPTASRAGCALS	1249
hCLASP1	-----TVNHADSRASLASLDSNPSTNEKSSEKTDNCEKI	PRPLALIGSTLRFDRLDQAE	1327
	..	..	
hCLASP4	IRSLLMCYLYIVKMISED	TLTYWNKVSPQELINILILLEVCLFHFYRMGKRNIARVHDA	1273
hCLASP5	TRNLMICFLWIMKNADQSLIRKWIADLPSTQLNRILDLLFICVLCFEYK	KGKQSSDKVSTQ	1268
hCLASP3	SRSLLICLLWVLKNADETVLQKWFTDLSVLQLNRLLDLLYLVCVSCFEYK	KGKVFERN	1384
hCLASP2	IKSLLMCFLYILKSMSDDALFTYWNKASTSELM	DDFTISEVCLHQFQYMGKRYIARNQEG	1287
hCLASP7	SRTLLACVLWVLKNT	EPALLQRWATDLTLPQLGRLLDLLYLCLAAFEYKGGKA	1309
hCLASP1	TRSLLMCFLHIMKTISYETLIAYWQ	RAPSPEVSDFFSILDVCLQNFYRLGKRNIIRKIAA	1387
	..	.....	
hCLASP4	WLSKHFGIDR-----KSQTMPALRNRS	GVMQARLQHLSSLESS-----	1311
hCLASP5	VLQKSRDVKAR-----LEEALLRGE	GARGEMRRRAPGNDRFPGLNEN---	1311
hCLASP3	TFKKS	KDMRAK-----LEEAILGSIGARQEMVRRSRGQLERSPSGSAFGSQ	1430
hCLASP2	LGPIVHDRKS-----QTL	PVSRNRTGMMHARLQQLGSLDNS-----	1323
hCLASP7	TFKKS	LDMKAR-----LEEAILGTIGARQEMVRRSRERSPFGNPEN-----	1350
hCLASP1	AFKFVQSTQNNGT	LKGSNPSCQTSGLLAQWMHSTSRHEGHKQHR	1442
		..	
hCLASP4	-----FTLNHSSTTTEADIFHQALLEGNTATEVSLTVLDTISFF	TQCFKTQLL	1359
hCLASP5	--LRWKKEQTHWRQANEKLDKTKAELDQ	EALISGNLATEAHLIILDMQENIIQASS-ALD	1368
hCLASP3	ENLRWRKDMTHWRQNT	EKLDKSRAEIEHEALIDGNLATEANLIILDTLEIVQTVS-VTE	1489
hCLASP2	-----LTFNHSYGHSDADVLHQSLLEANI	ATEVCLTALDTLSLFTLAFKNQLL	1371
hCLASP7	--VRWRKSVTHWKQTS	DRVDKTKDEMEHEALVEGNLATEASLVVLDTLEIIVQTVM-LSE	1407
hCLASP1	--ALS	NPKLLQMLDNTMTSNEIDIVHHVDTEANIEGCLTILDVSLFTQTHQRQLQ	1500
		.....	

FIG. 13  
4 of 6

hCLASP4 NNDGHNPLMKKVFDIHLAFLKNGQSEVSLKHVFASLRAFISKFPSAFFKGRVNMCAAFCY 1419  
hCLASP5 CKDS---LLGGVLRVLVNSLNCDSSTTYLTHCFATLRALIAKFGDLLFEEVEQCDFDLCH 1425  
hCLASP3 SKES---ILGGVLKVLHSMACNQSAVYLQHCFAQALVSKFPELLFEEETEQCADLCL 1546  
hCLASP2 ADHGHNPLMKKVFDVYLCFLQKHQSETALKNVFTALRS LIYKFPSTFYEGRADMCAALCY 1431  
hCLASP7 ARES---VLGAVLKVVLYSLGSAQSALFLQHGLATQALVSKFPELLFEEDTELCADLCL 1464  
hCLASP1 QCDCQNSLMKRGFDTYMLFFQVNQSATALKHVFASLRLFFVCKFPSAFFQGPADLCGSFCY 1560  
. : : : : \* \* : : : \* : : \* : : : \*

hCLASP4 EVLKCCTSKISSTRNEASALLYLLMRNFEYTKRKTFLRTHLQIIIAVSQLIADVALSGG 1479  
hCLASP5 QVLHHCSSSMDVTRSQCATLYLLMR--FSFGATSNFARVKMQVTMSLASLVGRAPDFNE 1483  
hCLASP3 RLLRHCS SIGTIRSHPSASLYLLMR--QNFEIGNNFARVKMQVPSLSSLVGTSQNFNE 1604  
hCLASP2 EILKCCNSKLSSIRTEASQLLYFLMRNFDYTGKKS FVRTHLQVIISVSQLIADVVGIGE 1491  
hCLASP7 RLLRHCGSRISTIRTHASASLYLLMR--QNFEIGHNFARVKMQVTMSLSSLVGTTQNFSE 1522  
hCLASP1 EVLKCCNHRSRSTQTEASALLYLFMRKNFEFNKQKSIVRSHLQLIKAVS QLIADAG-IGG 1619  
.:\*: \* : : : \* : : : \* : : : \* : : : \*

hCLASP4 SRFOESLFIINNFANS DRPMKATAFP AEVKDLTKRIRTVLMATAQMKEHEKDPEMLIDLQ 1539  
hCLASP5 EHLRRSLRTILAYSEEDTAMQMPFPTQVEELL CNLNSILYDTVKMREFQEDPEMLMDLM 1543  
hCLASP3 EFLRRSLKTI LTYAEEDLELRETTFPDQVQDLVFNLMILSDTVKMKEHQEDPEMLIDLQ 1664  
hCLASP2 TRFQQSLSIINNCANS DRLIKHTSFSSDVKDLTKRIRTVLMATAQMKEHENDPEMLVDLQ 1551  
hCLASP7 EHLRRSLKTI LTYAEEDMGLRDSTFAEQVQDLMFNLMILTDTVKMKEHQEDPEMLIDLQ 1582  
hCLASP1 SRFQHSLAITNNFANGDKQMKNSNFPAEVKDLTKRIRTVLMATAQMKEHEKDPEMLVDLQ 1679  
.:\*\* : : \* : : : \* : : : \* : : : \* : : : \*

transmembrane  
hCLASP4 YSLAKSYASTPELRK TWLDSMAKIHVKNGD FSEAAMCYVHVAALVAEFLHRKK----- 1592  
hCLASP5 YRIAKSYQASPDRLR TLWLQNMAEKHTKKKQYTEAAMCLVHAAALVAEYLSMLEDH----- 1598  
hCLASP3 YRIAKGYQTSPE-RLTWLQNMAGKHSERSN HAEAAQCLVHSAALVAEYLSMLEDR----- 1718  
hCLASP2 YSLAKSYASTPELRK TWLDSMARIHVKNGL LSEAAMCYVHV TALVAEYLTRKG----- 1604  
hCLASP7 YRIARGYQGS PDRLR TLWLQNMAAGKHAELGNHAEAAQCMVHAAALVAEYIALLEDQ----- 1637  
hCLASP1 YSLANSYASTPELR RTWLESMAKIHARNGD LSEAAMCYIHIAALIAEYLRKGYWKVEKI 1739  
\* : \* . \* : : \* \* : : \* : : \* : : \* : : \*

hCLASP4 -----LFPNGCSAFKKITPNIDEEGAMKEDAGMMD----- 1622  
hCLASP5 -----SYLPVGSVSFQNISSNVLEESVSEDTLS PDEDGV 1633  
hCLASP3 -----KYLPGCVTFQNISSNVLEESAVSDDVVS PDEEGI 1753  
hCLASP2 -----VFRQGCTAFRVITPNIDEEASMMEDVGMQD----- 1634  
hCLASP7 -----RHLPVGCVSFQNISSNVLEESAISDDILSPDEEGF 1672  
hCLASP1 CTASLLSEDT HPCDSNSLLTTPSGGSMFSMGWPAFLSITPNIKEEGA AKEDSGMHD----- 1795  
: \* : \* : : \* : : : \*

ITAM  
hCLASP4 ---VHYSEEV LLELLEQCVDGLWKAERYEI ISEISK LIVPIY EKRRFEKLTQVYRTIHG 1679  
hCLASP5 CAGQYFTESGLVGLLEQA AELFSTGGLYETVNEVYKLVIP ILEAHREFRKLTLTHSKLQR 1693  
hCLASP3 CSGKYFTESGLVGLLEQA AASFMSAGMYEAVNEVYKLVIP IHEANRDAKKLSTIHGKLQE 1813  
hCLASP2 ---VHFNE DVLME LLEQCADGLWKAERYELIADIYKLI IPIY EKRR----- 1677  
hCLASP7 CSGKHFTELGLVGLLEQA AGYFTMGGLYEAVNEVYKNLIP ILEAHRDYKKLA AVHGKLQE 1732  
hCLASP1 ---TPYNE NILVEQLYMCGEFLWK SERYELIADV NKPI IAVFEKQRDFKKLS DLYYDIHR 1852  
.:\* \*: \* : : \* : : : \*

ITAM DOCK motif DOCK motif ITAM  
hCLASP4 AYTKILEVMHTKKRLLGTFFRVAEY GQSF FEEEDGKEYIYKEFKLTGLSEISLR LVKIYG 1739  
hCLASP5 AFDSIVNKDH--KRMFGTYFRVGF FG-SKFGDLDEQE FVYKEFAITKLPEISHRLEAFYG 1750  
hCLASP3 AFSKIVHQSTGWERMFGTYFRVGF FG-TKFGDLDEQE FVYKEFAITKLAEISHRLEGEYG 1872  
hCLASP2 -----DFFEDEDGKEYIYKEFKLTPLSEISQRLLK IYS 1710  
hCLASP7 AFTKIMHQSSGWERVFGTYFRVGF FG-AHFGDLDEQE FVYKEFSITKLAEISHRLEEFYT 1791  
hCLASP1 SYLKVAEVVNSEKRLFGRYRVAFY GQG FFE EEGKEYIYKEFKLTGLSEISQRLLK IYA 1912  
\* : : : : \* : : \* : : \*

FIG. 13  
5 of 6

	ITAM	ITAM	
hCLASP4	EKFGTENVKIIQDS	DKVNAKELDPHYAHIQV	TYVKEYFDDKELTERKTEFERNHNISRFV 1799
hCLASP5	QCFGAEFVEVIKDS	TPVDKTKLDPNKAYIQI	TFVEEYFDEYEMKDRVITYFEKNFNLRRFM 1810
hCLASP3	ERFGEDVVEVIKDS	NPVDKCKLDPNKAYIQI	TYVEEYFDYEMKDRITYFDKNYNLRRFM 1932
hCLASP2	DKFGSENVKMIQDS	GKVNPKDLDSKYAYIQV	THVIEFFDEKELQERKTEFERSHNIRRFM 1770
hCLASP7	ERFGDDVVEIIKDS	YPVDKSKLDSQKAYIQI	TYVEEYFDYELKDRVITYFDRNYGLRTFL 1851
hCLASP1	DKFGADNVKIIQDS	NKVNPKDLDPHYAYIQV	TYVTEFFEEKEIEDRKTDfEMHHNINRFV 1972
	: ** : *::** : *	: ** : *::** : *	: ** : *::** : *
		ITAM	DOCK motif
hCLASP4	FEAPYTLGKKQGCIEEQCKRRTILTTSNSF	RYVKKRIPINCEQQIN	LKPIDGATDEIKD 1859
hCLASP5	YTTPFTLEGRPRGELHEQYRRNTVLTTMHAFF	YIKTRISVIQKEEFV	LTPIEVAIEDMKK 1870
hCLASP3	YCTPFTLDGRAHGEHLEQFKRKTILTTS	HAFFYIKTRVNVTHKEEI	LTPIEVAIEDMQK 1992
hCLASP2	FEMPFTQTGKRQGGVEEQCKRRTILTAIHCF	RYVKKRIPVMYQHHTD	LNPIEVAIDEMSK 1830
hCLASP7	FCTPFTPDGRAHGEHLEQHKRKTLLSTDHAF	YIKTRIRVCHREETV	LTPVEVAIEDMQK 1911
hCLASP1	FETPFTLSGKKHGGVAEQCKRRTILTTS	HLFFYVKKRIQVISQSST	LNPIEVAIDEMSR 2032
	: ** : * : * : *	: ** : * : *	: ** : * : *
	Coiled-coil		
hCLASP4	KTAELOKLCSSSTDVDMIQQLKLQ	QWVSVQVNAGPLAYARAF	LNDSSQASKYPPKKVSELK 1919
hCLASP5	KTLOLAVAINQEP	PPDAKMLQMVLOQ	SVGATVNQGP
hCLASP3	KTQELAFATHQDPADPKMLQMVLOQ	SVGTTVNQGP	LEVAQVFLSEIPSDPKLFRHHNKLR 2052
hCLASP2	KVAELRQLC	SSAEVDMIKLQKLQ	SVSVQVNAGPLAYARAF
hCLASP7	KTRELAFATEQDP	PPDAKMLQMVLOQ	SVGPTVNQGP
hCLASP1	KVSELNQLCTMEEVDMISLQKLQ	SVSVKVNAGPMAYARAF	LEETNAKKYPDQVKKLLK 2092
	* : *	* : *	* : *
	Coiled-coil		
hCLASP4	DMFRKFIQACS	IALELNERLIKEDQVEYHEGLKSNFRDMVKELSD	IIHEQILQEDTMHSP 1979
hCLASP5	LCFKEFIMRCGE	AVEKNKRLITADQREYQ	QELKKNYNKLKENLRPMIERKIPELYKPIFR 1990
hCLASP3	LCFKDFTKRCE	DALRKNKSLIGPVQKEYQRELGLSSP	----- 2090
hCLASP2	EVFRQFVEACG	QALAVNERLIKEDQLEYQEEMKANYREMAKELSE	IMHEQICPLEEKTS- 1949
hCLASP7	LCFKDFCKKCE	DALRKNKALIGPDQKEYHRELERNYCRLREALQPI	LTQRLPQLMAPTP- 2030
hCLASP1	EIFRQFADACG	QALDVNERLIKEDQLEYQEELRSHYKDMLSELSTVMNEQITGRDDLSKR 2152	
	* : *	* : *	* : *
	PDZ ligand		
hCLASP4	WMSNTLHVFC	AI	SGTSSDRGYSPRYAEV-- 2008
hCLASP5	VESQKRDS	FHRSSFRKCETQLSQGS	----- 2015
hCLASP3	-----		
hCLASP2	VLPNSLHIF	NAISGTPTSTMVHGMTSS	SSVV 1980
hCLASP7	--PGLRNSLN	RASFRKADL	----- 2047
hCLASP1	GVDQTCTRV	ISKATPALPTVSISS	SAEV-- 2180

FIG. 13  
6 of 6

# Human CLASP-2 expression in T cells upon activation

**A)**

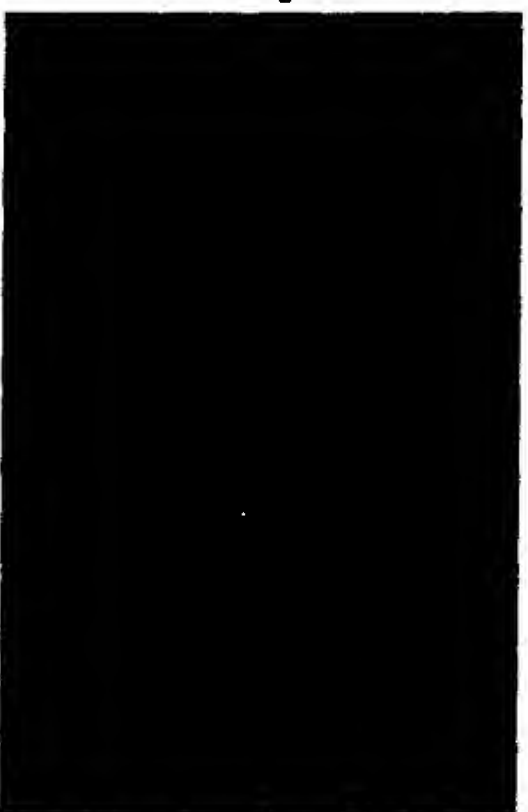
hours post  
activation

0 1 2 4 8 14

~ 7.5 kb -

CLASP-2

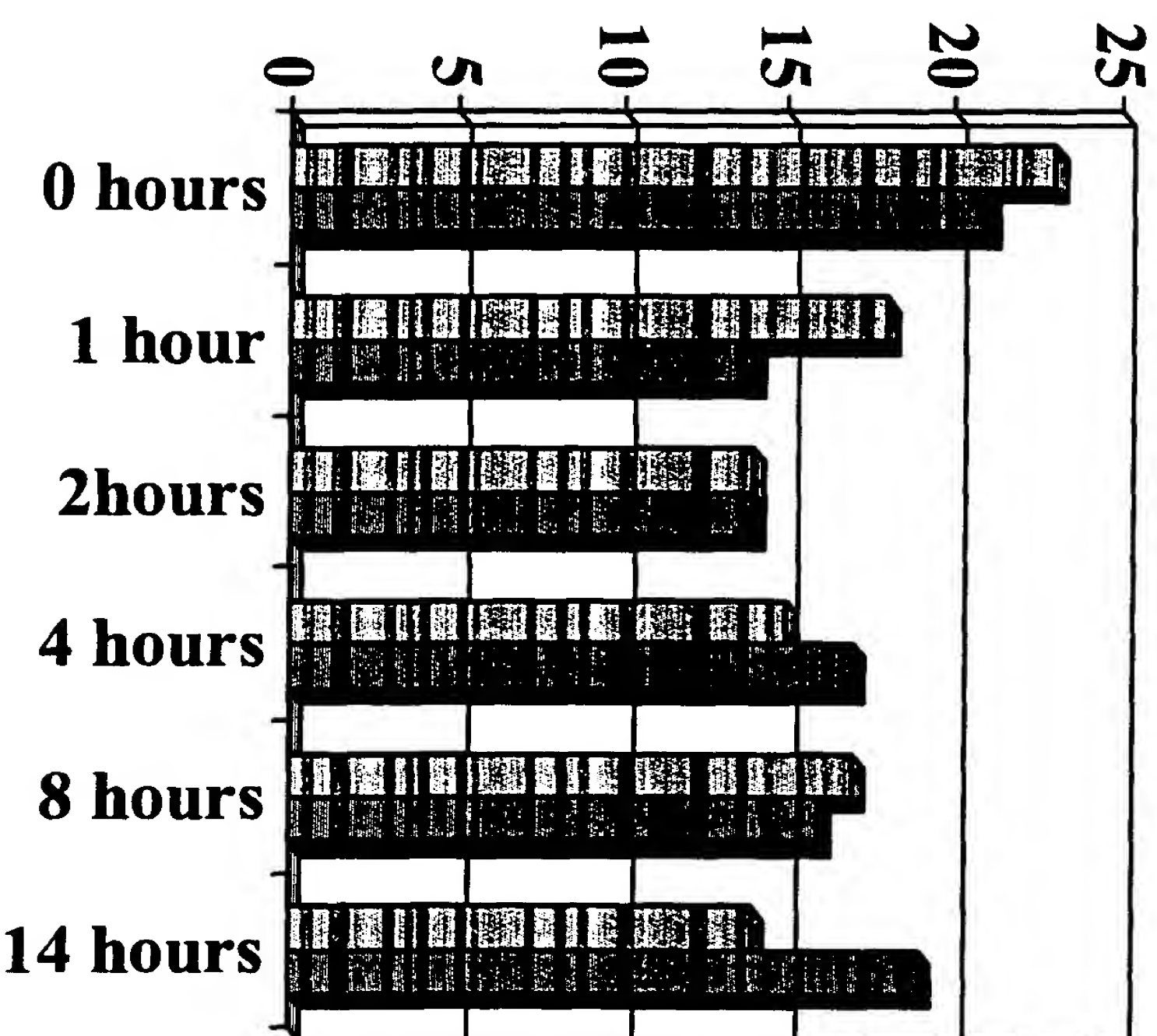
2.5 kb -



28s rRNA Ethbr. staining

**B)**

% of total  
signal



CLASP-2 (A)  
CLASP-2 (B)

hours post  
activation

09687837 101300